Distribution Agreement

Three Essays in Environmental and Resource Economics

By

Stephen F. Kiebzak, III
Doctor of Philosophy

Economics

Paul H. Rubin, Ph.D.
Advisor

David Jacho-Chavez, Ph.D.
Committee Member

Tilman Klumpp, Ph.D.
Committee Member

Kurt E. Schnier, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
James T. Laney School of Graduate Studies

Date

Three Essays in Environmental and Resource Economics

By

Stephen F. Kiebzak, III
M.A., The University of Alabama in Huntsville, 2006

Advisor: Paul H. Rubin, Ph.D., Purdue University

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2012

Abstract

Three Essays in Environmental and Resource Economics
By Stephen F. Kiebzak, III

This dissertation studies three important and timely topics related to the use of natural resources and the generation of pollutants. Chapter One explores the impact that state water management frameworks have on farm productivity. Prior appropriation law states allow greater access to surface waters for use in irrigation, a critical input for crops. I find that states operating under this form of water law have corn yields that are approximately 20 to 30 bushels per acre higher than their riparian law counterparts. This represents an 18 to 28% increase on average. Chapter Two addresses to what extent oil producers respond to changes in price and whether higher royalties on oil production result in a reduction in the life of a producing lease. By exploiting a unique lease-level data set of monthly sales of oil from leases on federal properties, I estimate that production from the vast majority of currently producing leases is highly inelastic. Estimated elasticities are small and generally not significantly different than zero. This data set also includes data from the fourteen-year period during which marginal leases were granted royalty reductions by the Bureau of Land Management to stimulate production during periods of low price. The most marginal of this class of leases, those that do not report sales regularly, have significantly higher elasticities. Further, leases that participated in the royalty reduction program had a 15% lower probability of being shut-in than those leases that were not eligible for the program. Finally, Chapter Three investigates a novel method to predict carbon dioxide emissions from developing countries, the primary driver of emissions growth over the past decade. I employ an environmental Kuznets curve-type analysis to predict emissions, but rather than relating the level of per capita pollutant to a country's gross domestic product, I use a socio-economic status measure constructed from household characteristics and possessions survey data from developing countries. This approach improves on in-sample prediction of emissions which rely on gross domestic product alone, although data limitations prevent formal testing of this conclusion.

Three Essays in Environmental and Resource Economics

By

Stephen F. Kiebzak, III
M.A., The University of Alabama in Huntsville, 2006

Advisor: Paul H. Rubin, Ph.D., Purdue University

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2012

Acknowledgements

This dissertation would not have been possible without the generous support of numerous individuals. I would like to thank my advisor, Paul Rubin, for the support and guidance he provided during this process. I am also truly indebted and thankful to my committee members for their encouraging words and thoughtful criticism. Tilman Klumpp was incredibly generous with his time and advice in writing this dissertation and completing this process successfully. David Jaco-Chavez provided critical technical advice and brought a depth of econometric knowledge to bear in assisting me. Kurt Schnier provided excellent research advice and helped to inspire my interests in Environmental and Resource Economics. I owe sincere thanks to my classmates Griffin Edwards, Mzwandile Ginindza, Erik Nesson, and Josh Robinson for their encouragement and advice on issues large and small related to the completion of this work. Their five years of friendship was crucial in completing this dissertation. I am also obliged to Andrew Stocking and Tim Fitzgerald for their wise and helpful advice. Further, Maria Arbatskaya, Sean Mann, Kimberly Lorch, Marie Browne, Alexa DeVetter, and Brenda Baxter made this process much easier with their kind words and excellent logistical help. Finally, I owe a debt of gratitude to my friends and family. Their unconditional love and encouragement provided a strong base of support as I worked to complete this dissertation.

Contents

List of Figures

List of Tables

1 An Economic Analysis of The Role of Water Law in Improving Corn Yields

Abstract

This paper studies the effect of U.S. state water laws on irrigation rates and corn yields. Adequate supply of water is important for crop development and has been shown to impact the amount of crop produced per acre of land. Rights governing the use of water therefore play a role in farm productivity. The increased access to surface water resources allowed in a prior appropriation system provides an opportunity to evaluate the potential for yield improvements in riparian states if they were to allow non-riparian farms similar access to surface water. I estimate that allowing non-riparians access to surface water for irrigation increases corn yields in prior appropriation states by between 20 and 30 bushels per acre on average over states with a riparian water law system, an 18 to 28% increase. For an individual non-riparian farm in a riparian water law state, this would amount to an increase in corn revenues of between \$80 and \$180 per acre, depending on corn prices.

1.1 Introduction

Access to water at critical times before and during the growing season for plant evapotranspiration, whether through precipitation or irrigation, is important to the feasibility and profitability of row crop farming. As westward expansion of the United States began in earnest during the latter part of the nineteenth century, it became apparent that eastern water laws were inadequate for supplying sufficient amounts in the semi-arid west (see for example, *Cadillac Desert* by Reisner 1993). The common law riparian system of the eastern states associates rights to make “reasonable” use of surface waters with the ownership of land adjacent to the water source (Beck 2004). In the wet eastern states, this legal structure still allows widespread “dryland farming” on non-riparian lands. When needed, supplemental irrigation can be supplied by groundwater wells in areas where access to sufficient groundwater supplies is not restricted, although the costs associated with drilling wells and pumping water can be prohibitive.

To cope with the need for additional water for irrigation and other uses in the dryer states west of the 100th meridian (e.g. California gold mining), laws evolved to disassociate rights to use water from the ownership of adjacent land requirement. Instead, western laws began to link the right to use water for "beneficial" purposes to the timing of the initial withdrawal and use. In this prior appropriation system, the first ones to apply the water to broadly defined beneficial uses established rights to that quantity of water, generally regardless of the location of the use in relation to the source water. In times of drought the more senior appropriators of water, those who established rights earlier, can continue their withdrawals while the most junior appropriators must cease withdrawals. Importantly, unlike in riparian law states, in prior appropriation states rights to water can be bought, sold, or otherwise transferred to help direct resources to higher value uses. Figure 1 shows the breakdown of states by water type.[1]

This paper considers how the main difference between these two types of water laws, the possible use of surface water by non-riparians in the west, impacts agricultural outcomes. I develop a model to estimate the impact of water law on irrigation rates and in turn how irrigation rates impact corn yields. Corn was chosen for three reasons. First, there is widespread cultivation of various types of maize throughout the United States, allowing a large pool of counties to draw on. Second, the weather, technology, and crop yield relationship at the core of the model has been frequently applied to corn production (Smith 1914; Wallace 1920; Thompson 1968, 1975, 1986, 1988; Garcia et al. 1987; Tannura et al. 2008), although previous studies focus on primarily eastern states without accounting for the impact of irrigation. Third, U.S. corn production accounts for approximately 42% of total global production,[2] and any major impacts of state policies on corn yields could potentially have large global implications.

[1] U.S. state water law, while modeled here as either prior appropriation or riparian, is actually more diverse. There are variations on both types of systems, however, the key difference remains: the ability of non-riparians in the west to establish rights to use surface waters for irrigation. While several eastern states have moved towards more regulated water management frameworks, these regulations don't modify the underlying difference exploited by this study. Beck (2004) provides a detailed account of each states' legal structure relating to water use. Florida has a legal approach to water use which can not be readily categorized as either prior appropriation or riparian for the purposes of this analysis and it was therefore excluded.

[2] USDA Foreign Agricultural Service, `http://www.fas.usda.gov/psdonline/psdHome.aspx`

I compile a panel of county-level corn yield, precipitation, temperature, farm characteristics, and irrigation data for 1995, 2000, and 2005. I then filter the data set to include only those counties where corn farming comprises at least 70% of harvested cropland. The resulting unbalanced panel is constrained in this fashion due to limitations of the irrigation data which are not broken out by crop type, but rather county totals for the year, and are only compiled every five years.

Using panel regression, as well as estimation of a two-stage model with pooled data, I estimate that allowing non-riparians access to surface water increases corn yields in prior appropriation counties by approximately 28 bushels per acre on average over counties in riparian water law states. These results are robust to higher thresholds of corn in the county, time trending specifications, exclusion of the counties in Nebraska (the main corn producing state operating under a prior appropriation legal framework) and exclusion of counties in southwestern states that may have multiple plantings in a year. Using these alternate specifications, corn yields in counties in prior appropriation states are consistently between 20 and 30 bushels per acre higher than in counties in riparian law states. For an individual, non-riparian farm in a riparian water law state, this range of yield enhancements would amount to an increase in revenue from corn of between $80 and $180 per acre[3].

The remainder of the paper is organized as follows. Section 2 provides a brief overview of some of the literature on water rights and summarizes the literature on modeling crop production, focusing primarily on those studies related to corn yields. Sections 3 and 4 describe the empirical analyses employed in the study and the data used, respectively. Section 5 presents the results from regression, testing and robustness checks, and Section 6 concludes.

1.2 Literature Review

Much has been written about the common law riparian legal constraints on water use in the eastern U.S. from the legal and policy perspective (see for example, Maloney et al.

[3]These figures were calculated for corn grain prices in the range of $4 to $6 per bushel and do not deduct the expenses of getting the water from the source to the field or the irrigation equipment necessary to apply the water.

1968; Eheart 2002; Dellapenna 2004; Marcus and Kiebzak 2008; Klein et al. 2009). Florida, finding the nature of rights to water usage under a riparian system inadequate, went so far as to convert to a statutory system based on permitting designed by legal experts at the University of Florida (Maloney et al. 1972). Marcus and Kiebzak (2008) and McNider et al. (2005) note that despite the semi-arid and desert climates of the western states, they often have higher returns from agriculture than very wet eastern states such as Alabama. Even in 2000, one of the driest years on record for Alabama, the flow of the Alabama river in Monroe County was still a massive 10 million acre-feet per year. Yet farm crop receipts were lower in Alabama than in New Mexico where the lower Rio Grande, the primary source of surface water irrigation, averages a paltry 790,000 acre-feet per year released flow (McNider et al. 2005). Marcus and Kiebzak conclude that water law, and not the availability of water resources, is the primary impediment to improving yields, and therefore agricultural revenues. Given reduced rainfall in the west and legal constraints to surface water uses for irrigation in the east, it is no surprise that the 17 western states use roughly 88% of the total amount of water used for irrigation in the United States annually (Schlenker et al. 2006).

Crop yields in general, and corn yields in particular, have been studied in many contexts. One of the primary areas of research has been the modeling of corn yield distributions for crop insurance purposes (Ramirez 1997; Goodwin and Ker 1998; Ramirez et al. 2003; Norwood et al. 2004; Ozaki et al. 2008; Harri et al. 2009). The University of Missouri's Food and Agricultural Policy Research Institute reports that over 80% of the nation's corn acres participated in insurance programs in 2009. This amounted to more than 70 million acres of corn covered by insurance, approximately 10 million more acres than soybeans, the next highest. Further, over half of the insured acres had coverage levels of 70% or higher (FAPRI 2010). In 2008/2009 total indemnities received by producers of all crops were more than $8 billion. In light of the large financial incentives involved in setting appropriate premiums, the volume of research on corn yield distribution modeling is not surprising.

Several agronomists have approached crop yields from the standpoint of the interactions between soil, nutrients, and weather through the development of complex growth simulation models (Jones and Kiniry 1986; Duchon 1986; Kaufmann and Snell 1997). These models are founded on biology and experimental data from test plots. However, obtaining the detailed information needed for the specific model parameters becomes problematic when one proceeds beyond the plot level and attempts to apply the models more broadly to large geographic areas, such as counties or regions. Hook (1994) provides one example of the limited scope of such models. He uses crop growth and water use models for corn, peanuts, and soybeans to estimate irrigation needs in the coastal plains region of south Georgia to prevent crop losses during periods of drought. These models have also been used to optimize the use of limited water resources. Paudel et al. (2005) combine crop simulation models with hydrological and dynamic economic models to aid in irrigation water allocations among different crops, including corn, during periods of water shortage in the southeast.

A third area of corn yield research focuses on using regression models to estimate the impacts of weather and technology on crop growth. These studies are aimed primarily at determining what patterns of rainfall and temperature are optimum for the growth of corn in the eastern U.S. (Runge and Odell 1958; Thompson 1968, 1975, 1986, 1988; Baier 1977; Swanson and Nyankori 1979; Garcia et al. 1987; Dixon et al. 1994; Andresen et al. 2001; Hu and Buyanovsky 2003; Tannura et al. 2008). They generally consider aggregate yields for large geographic areas—usually states—based on measures of average state precipitation totals and temperatures.[4] These authors consider farms within a limited region around the U.S. Corn Belt consisting of only riparian states. Implicit in these analyses is the assumption of similar patterns and rates of irrigation. Further, average farm characteristics are ignored.

Finally, Mendelsohn and Dinar (2003), expanding the model of Mendelsohn et al. (1994), attempt to measure the monetary impacts of irrigation through a so-called Ricardian approach that estimates how farm values are linked to climate, soil, and water. While they incorporate data on irrigation rates and technology, their dependent variable of interest is the value of farms from the 1997 U.S. Census of Agriculture. As they admit, the values

[4]Exceptions are Runge and Odell (1958), who study how local precipitation and temperature measurements impact corn yields on an experimental farm, and Hu and Buyanovsky (2003), who study yields over a 100 year period in central Missouri.

are self-reported and based only on the farmers' estimates of the market values of their land and buildings. They find that the value of irrigated farms is not sensitive to precipitation, as would be expected. Schlenker et al. (2006) use the same Ricardian approach, but instead incorporate climate variation with an alternative specification based on the concept of growing degree days.[5] Schlenker and Roberts (2006) model a nonlinear relationship between a modified measure of growing degree days and corn yield. Their modified degree days measure uses daily temperature data from 2.5 by 2.5 mile grids over the eastern U.S. combined with data on crop locations within these grids. This research was aimed at assessing the impact of temperatures on yield to aid in yield predictions in a warming climate. Schlenker and Roberts only apply their model to riparian states and assume relatively constant patterns of irrigation.

Despite this body of work, a gap exists between the research that considers yields and the studies looking at the impact of irrigation on farm values. The present study bridges this gap, at least partially, by estimating how irrigation has impacted the actual yields of a particular crop over a large geographic area. Further, previous agricultural economics research has ignored the impact of legal frameworks relating to water use. This analysis, therefore, breaks new ground in that dimension.

1.3 Empirical Analysis

In a prior appropriation system the initial right, and therefore allocation of the water, is established by who uses it first. Importantly, as the rights can be bought, sold, or traded, in theory water will be allocated to where its marginal product is highest. Most prior appropriation states do place geographic limits on trades (within the same river basin, for example), and to what use the traded water can be put (irrigation water rights sales may be restricted to purchases for irrigation). However, the direction of trade will be toward higher marginal product of the water, unlike the static water right in a riparian law state

[5] As they explain, growing degree days are defined as "the sum of degrees above a lower baseline and below an upper threshold during the growing season...a day with a temperature below $8°C$ results in 0 degree days; a day with a temperature between 8 and $32°C$ contributes the number of degrees above $8°C$, and a day with a temperature above $32°C$ contributes 24 degree days. Degree days are then summed over the growing period".

which is tied to the ownership of land. While land can be traded or merged or used for other purposes, this is a cumbersome way to transfer water rights. Further, a given amount of riparian land can only make productive use of a certain amount of water, particularly for agriculture. After a certain level of irrigation, additional water can hurt crop development. Additional value can then only be realized by using the water on lands which may not be held by the riparian landowner.

Although we are interested in testing the higher relative efficiency of water allocation under a prior appropriation system, we cannot observe the marginal product of water directly. Measuring the efficiency gained in a prior appropriation system can, however, be accomplished using a predominant agricultural measure. The productivity of farms are measured by the average product of land, defined as yield. In the case of corn, this would be the number of bushels produced on the farm per acre of land planted in corn. Water application increases plant production by increasing the moisture content of the soil where it can then be drawn into the plant along with the nutrients it extracts from the soil. Water is applied over a fixed unit of land—an acre—at a particular flow rate. Also, water volumes are frequently measured using a unit of land as a reference, in acre-feet—the volume of water that would correspond to that many feet of water sitting on an acre of land. Using the average product of land to convey productivity improvements is, therefore, a convenience which will provide a similar result should the average product of water be measured directly. Further, using yield as a dependent variable will allow for consistency and comparability with other research on agricultural outcomes.

I expand the models proposed by Thompson (1968) and Tannura et al. (2008) to incorporate the effect of water law on irrigation rates, and irrigation rates on yields. Instead of aggregating the data to state level averages, I get richer variation by using county level climate and yield data. This model also expands on previous research by incorporating average county farm characteristics.[6] As the type of water law directly impacts the access

[6] I currently only use average farm size, but will extend this to include potentially available data on chemical usage and average soil quality, as represented by the soil's K-factor. While farm level data accounting for plant populations, maize hybrid type, grade of product, and soil characteristics would be ideal, such a data set of any scale is not presently available to researchers. Regressions using percentage of county employed in agriculture and percentage of the population who are immigrants were used, but these regressors were not significant and did not appreciably impact the results.

by non-riparians to irrigation water, and hence irrigation rates, but should not directly impact yields, a two-stage model is appropriate:

$$Y_{it} = \alpha_1 + \gamma_{1t} + \beta I_{it} + \boldsymbol{\zeta}_{\mathbf{1}}^{'}\mathbf{F_{it}} + \boldsymbol{\delta}_{\mathbf{1}}^{'}\mathbf{X_{it}} + \epsilon_{it}, \tag{1}$$

$$I_{it} = \alpha_2 + \gamma_{2t} + \eta L_i + \boldsymbol{\zeta}_{\mathbf{2}}^{'}\mathbf{F_{it}} + \boldsymbol{\delta}_{\mathbf{2}}^{'}\mathbf{X_{it}} + \upsilon_{it}. \tag{2}$$

Y_{it} is the corn yield per acre in county i in year t; I_{it} is the average daily rate of irrigation per acre; γ_{1t} and γ_{2t} are year effects, which capture relevant factors that are roughly equivalent across counties, such as changes in technology related to irrigation or tillage and seed hybrids, for example; L_i is an indicator of water law type in county i, taking on values of 1 for counties in prior appropriation states and 0 for counties in riparian law states; F_{it} is a vector of relevant average farm and county characteristics (see note 6); X_{it} is a vector of average monthly temperatures and their squares, total monthly precipitation amounts and their squares, and interactions between monthly temperature and precipitation for the predominant months of the corn growing season (April - September); and υ_{it} and ϵ_{it} are contemporaneous error terms. Water laws might be considered an endogenous treatment in an irrigation-yield model since the adoption of prior appropriation law was based on the need for additional water in the dry western states. However, by controlling for climate with the exogenous temperature and rainfall measurements, the outcomes and treatment are conditionally independent in this situation.

One approach is to estimate the two-stage model using control function techniques (instrumental variable regression). To estimate the impact of water law, L_i, on yield, Y_{it}, through irrigation, I_{it}, our statistic of interest would then be $\hat{\eta} \times \hat{\beta}$. If the highly unbalanced panel described in Section 1.4 is poolable, the model can be estimated using two-stage least squares. Alternatively, the model can be estimated with the instrumental variable method developed by Hausman and Taylor (1981) for panel data with time-invariant regressors. However, given the triangular nature of the system described in (1) and (2), a convenient simplification can be made by substituting the irrigation equation, (2), for I_{it} in the yield equation, (1).

$$Y_{it} = (\alpha_1 + \beta\alpha_2) + (\gamma_{1t} + \beta\gamma_{2t}) + \beta\eta L_i + (\boldsymbol{\zeta}_1^{'} + \beta\boldsymbol{\zeta}_2^{'})F_{it} + (\boldsymbol{\delta}_1^{'} + \beta\boldsymbol{\delta}_2^{'})X_{it} + (\epsilon_{it} + \beta\upsilon_{it}). \quad (3)$$

Equation 3 can now be estimated using panel techniques. Panel generalized least squares is used with random and fixed effects estimators. While fixed effects estimation can not identify $\widehat{\beta\eta}$ as water law is time-invariant, a Hausman test does not reject the consistency and efficiency of the random effects estimator, as discussed in Section 1.5.1.

Estimation of the models, however, is complicated by the multicollinearity of the individual climate terms. Multicollinearity is also present, although with smaller correlations, between irrigation rates and water law, and the climate terms. As a result, coefficient estimates may be imprecise with large standard errors. My interest, however, is not in determining the effect of the individual climate variables, but rather to use as controls a set of measures that characterize climate, retaining the variation of the exogenous climate variables across regions and time. I therefore combine the climate variables using weightings from a principal component analysis (PCA). The components are created by applying a specific weighting to each of the climate variables as follows:

$$\begin{aligned} PC_1 &= a_{11}X_1 + a_{12}X_2 + \ldots + a_{1n}X_n, \\ &\vdots \\ PC_n &= a_{n1}X_1 + a_{n2}X_2 + \ldots + a_{nn}X_n. \end{aligned}$$

$PC_1,\ldots, PC_n$ is the vector of principal components and $X_1,\ldots,X_n$ is the vector of climate variables being combined (in this case the average monthly temperatures and their squares, total precipitation and their squares, and the interactions between temperature and precipitation). Finally, $a_{11},\ldots, a_{nn}$ is the matrix of weights associated with each principal component and variable. In a PCA using unstandardized data, the weights are the eigenvectors of the data's correlation matrix. Each principal component then accounts for a portion of the total variation in the original data. The amount of variation associated with each of the n principal components is calculated as the eigenvalue, λ, of each eigenvector divided by n. The components are ordered so that PC_1 explains the largest variation and PC_n explains the smallest.

I then replace the climate variables, X_{it}, in equations (1) and (2) with the principal components accounting for the largest percentages of variation in the original data. Multiple selection methods are tried. I use the Kaiser-Guttman Rule, only selecting those components with eigenvalues greater than 1.00, (Draper and Smith 1981; Loehlin 1998) and also apply the method in Myers (1986) by eliminating those components with the smallest t-statistics. The final two-stage model that results is shown in (4) and (5) below.

$$Y_{it} = \alpha_1 + \gamma_{1t} + \beta I_{it} + \boldsymbol{\zeta_1'}\mathbf{F_{it}} + \boldsymbol{\delta_1'}\mathbf{PC_{it}} + \epsilon_{it}, \tag{4}$$

$$I_{it} = \alpha_2 + \gamma_{2t} + \eta L_i + \boldsymbol{\zeta_2'}\mathbf{F_{it}} + \boldsymbol{\delta_2'}\mathbf{PC_{it}} + \upsilon_{it}, \tag{5}$$

or the simplified version,

$$Y_{it} = (\alpha_1 + \beta\alpha_2) + (\gamma_{1t} + \beta\gamma_{2t}) + \beta\eta L_i + (\boldsymbol{\zeta_1'} + \beta\boldsymbol{\zeta_2'})F_{it} + (\boldsymbol{\delta_1'} + \beta\boldsymbol{\delta_2'})PC_{it} + (\epsilon_{it} + \beta\upsilon_{it}). \tag{6}$$

If the estimated coefficients of the principal components are of interest, testing requires that their standard errors be adjusted by dividing each component's estimated standard error by the square root of the component's associated eigenvalue, λ (Fekedulegn et al. 2002).

1.4 Data

Water law information for each state was compiled from Beck's *Waters and Water Rights* (2004).[7] Corn yield data was obtained for all U.S. counties where corn was grown from 1985 through 2010 from the U.S. Department of Agriculture's National Agriculture Statistics Service. The yields are not broken out by maize type, but are rather aggregate bushels produced in the county per acre of harvested corn crop. Precipitation and temperature data were obtained from the National Oceanic and Atmospheric Administration's National

[7] in Marcus and Kiebzak (2008)

Climate Data Center. Precipitation is reported as total inches per month and temperature is reported as the average monthly value. This climate data is reported from each NOAA monitoring station. Where there is more than one monitoring station per county, the values are averaged to obtain a mean temperature and a mean precipitation total for each county each month.

The irrigation data was obtained from the U.S. Geological Survey's *Estimated Use of Water in the United States* reports. The USGS has estimated and compiled annual water use by county and type of use from 1985 through 2005 in five-year increments.[8] Unfortunately for the researcher, this data is not supplied by crop. Daily irrigation rates therefore include water used for all crops over an entire year period. If crops other than corn are selectively irrigated, irrigation rates will be overestimated. I therefore only consider counties where corn is the predominant harvested crop. To create this filter, I obtained data on the amount of each crop harvested per county from the USDA's Census of Agriculture. The census is conducted every five years going back to 1987. However, only data from the 1997 census onward is readily available.[9] As this data is not from the same years as the USGS data on irrigation rates, care must be taken with its use. Two methods can be used at this point to generate the data set. First, an assumption could be made that irrigation rates did not change significantly between, for example, 1995 and 1997, and the panel could be created for the years of the Census. Conversely, an assumption can be made that the relative amounts of corn compared to other crops in the county did not change drastically in the two-year interval between the collecting of irrigation data and the Census. The latter assumption seems more plausible, particularly over two-year intervals, and is not as likely to alter the results given that the percentage of corn in the county is only used as a filter. To check the validity of the filter, I perform robustness checks by estimating the model at minimum thresholds of corn as a percentage of all crops in the county between 50% and 90%.

[8]See the USGS *National Handbook of Recommended Methods of Water Data Acquisition*, available at `http://pubs.usgs.gov/chapter11/chapter11H.html`, for details on the exhaustive array of methods used to estimate irrigation water withdrawals, particularly where direct, metered data is not available.

[9]The data set used in the present study could be expanded with the USDA's assistance in obtaining data from the 1987 and 1992 censuses

County farm characteristics (see note 6), are also obtained from the Census of Agriculture. Farm sizes may not change significantly in the two-year intervals, however, I run the models with and without farm characteristics to see if they are of great import to the estimation results. The data on county employment in agriculture and immigrants as a percentage of county population were obtained from the USDA Economic Research Service's *Rural Atlas*.[10]

The resulting panel is highly unbalanced, consisting of 234 total observations over the three years for which data is readily available. 170 unique counties in 18 states are included, however, only 54 counties have repeated observations over time and only 11 counties have observations in each of the three years. Approximately 45% of the observations are in states with prior appropriation water laws. Table 1 summarizes the data for those counties with greater than 70% corn.

Given the aggregation of the irrigation data, summary statistics of corn yields in all counties in prior appropriation and riparian law states are misleading. If the generally wetter climate of the east and the use of supplemental groundwater irrigation were enough to offset the lack of access to surface water by non-riparians, we should not see any significant difference in yields between states operating under different water law frameworks. A higher average yield in the east may even result given the over appropriation of western waters. Indeed, by looking at summary statistics alone one might make this conclusion. Overall corn yields in riparian states for the years considered in this study averaged 121 bushels per acre, but only 113 in prior appropriation states. This, however, provides a misleading picture of the impact of the laws on corn yields. Given a generally fixed access to water, farmers are assumed to allocate their land and water resources among a variety of crops to produce the highest return from their produce[11]. Within a county and even within an individual farm, multiple crops are grown. Farmers may preferentially irrigate crops of higher value (Paudel et al. 2005), and they may allocate corn to less desirable plots or rely more on dryland farming for corn, where precipitation accounts for the only source of water.

[10] Available at `http://www.ers.usda.gov/data/ruralatlas/download.htm`

[11] see for example the Crop Water Allocator software developed at Kansas State University at `http://www.ksre.ksu.edu/mil/cwa/`, as well as Paudel et al. (2005)

If this is occurring in the west, one would expect the wetter climate of the eastern states to produce higher corn yields. Eastern farmers may also make this calculus and preferentially grow higher-valued crops on plots adjacent to surface water sources, but given the wetter conditions, higher corn yields should still be obtained.

When we consider those counties where corn is the primary crop, a different story emerges. In these counties, switching out of corn and into other crops has not occurred to any large degree for that particular observation year, and one should be able to isolate the effects of the different laws on yields if enough such counties could be identified operating under each legal system. While not sufficient to draw any conclusions, descriptive statistics for counties where corn accounts for at least 70% of harvested cropland show an 18 bushel per acre higher average yield in prior appropriation states compared to their wetter eastern neighbors as shown in Table 2 (yields of 128 in prior appropriation counties versus 110 bushels per acre in riparian counties with more than 70% corn).

1.5 Results

1.5.1 Regression using individual climate variables

Regression results of equations (1) and (2) pooling the data are displayed in the "2SLS" column of Table 3. I also estimate equation (3) using panel generalized least squares (GLS) with random effects, and with pooled data by ordinary least squares (OLS). Given the multicollinearity of the climate variables, few of them are significant (8 of 30 in the first stage and 5 of 30 in the second stage are significant at the 5% level). All climate variables are, however, jointly significant.

A bootstrap Hausman test with robust standard errors, comparing the two-stage model and a random effects estimation, fails to reject endogeneity of the irrigation rate. Further, the partial R^2 and F-test of the instrument indicate that water law is a good instrument for irrigation rate. A Hausman test also fails to reject the null of a random effects model versus a fixed effects alternative (the estimated density of the individual-specific effects is in Figure 2; fixed effects estimation results are available upon request). An F-test of poolability of the data over the three observation years, based on the residual sum of squares from OLS regressions, fails to reject poolability over time. However, the Breusch-Pagan Lagrangian multiplier test rejects an OLS error structure in favor of random effects. Wooldridge's (2002) test of serial correlation rejects the null of no first order autocorrelation at the 1% level. Therefore, cluster-robust standard errors clustered at the county level are used. Both pooled OLS and two-stage estimations with climate variables indicate approximately 46 bushels per acre higher corn yields in prior appropriation counties.[12] The random effects estimator is only slightly lower at 43 bushels per acre.

While the effect of average county farm size is significant in the random effects and OLS estimations, the magnitude is quite small (an increase of 100 acres in average county farm size decreases corn yield by about a third of a bushel). Finally, the year dummies show increased yields in 2000 and 2005 relative to 1995 in the random effects model, but they are only significant at the 10% level. A year trend in place of year dummies was also considered for comparison with Tannura et al. (2008). As expected, there was no significant difference observed in the impact of water law by accounting for technology changes with either year dummies or a time trend. The average annual yield increase of 2.9 bushels per acre estimated in the present study using a time trend variable is only slightly higher than that obtained by Tannura et al. for their smaller three-state study between 1960 and 2006.

1.5.2 Regression using principal components

Table 4 shows the correlation between the climate variables in the entire data set. As expected, precipitation rates for each month of the growing season are highly correlated, with a minimum value of 0.68 and average of 0.76. Monthly temperatures are even more highly correlated, with a minimum correlation of 0.73 and average of 0.82. While there is no defined threshold to consider multicollinearity severe (Goldberger 1991), the lack of significance of climate variables well-known to influence crop yields is an indicator that it is influencing regression results. To eliminate the correlation between climate variables, a set of principal components are generated that are, by construction, orthogonal. The climate

[12]In the case of the two-stage model, having a prior appropriation water law increases irrigation by about 1.9 thousand gallons per day per acre ($\hat{\eta}$) as estimated in the first stage, and an increase in irrigation of one thousand gallons per day per acre increases corn yields by 23.8 bushels per acre ($\hat{\beta}$) from the second stage. The impact of law on yield is therefore approximately 46 bushels per acre.

variable weights from the principal component analysis are displayed in Table 5 for the four components with eigenvalues greater than one. These four components collectively account for 86% of the variation in the original climate data.

Estimation using these four principal components in place of the climate variables markedly improves the regression results. Two-stage regression results of equations (4) and (5) are displayed in the "2SLS" column of Table 6. Again, this result is compared to estimation of equation (6) by GLS with random effects as well as pooled ordinary least squares. Results from testing are consistent with those from estimates using the individual climate variables as regressors. The only exception is that Wooldridge's (2002) test of serial correlation now fails to reject the null of no first order autocorrelation. Cluster-robust standard errors were still used and all coefficients remain highly significant. As in panel GLS estimation using climate variables, regression using principal components fails to reject random effects over a fixed effects model. Figure 2 compares the distribution of the individual-specific effects from the GLS fixed effects models using climate variables to that using principal components. While a Hausman test fails to reject the random effects model in either case, the model using principal components to account for climate variation removes the outlying individual-specific effects. Although in both cases the effects are approximately normally distributed.

Using principal components provides a more realistic estimate of the impact of water law on corn yields. The multicollinearity inherent using the individual climate data appears, therefore, to be inflating the impact of water law on yields. Both the pooled OLS and two-stage estimators show higher yields of approximately 27 bushels per acre in prior appropriation counties compared to counties in riparian states. The GLS random effects regression estimates 28 bushels per acre higher yield in prior appropriation counties. Two-stage estimation indicates that prior appropriation counties have, on average, 2.5 thousand gallons per day per acre higher irrigation rates than counties in riparian law states, controlling for farm size, year, and climate variables (through principal components). Farm size is now significant using all three estimators, as are all year dummies. While farm size is significant, the direct impact of farm size on yield is quite small. However, including farm size, as in regression using the climate variables, increases the estimated impact of irrigation on yield. Without controlling for farm size, the impact of water law on corn yield is reduced to approximately 22 bushels per acre higher yield in prior appropriation counties.

The first stage results, in which the first, second, and fourth principal components ($PC1$ and $PC4$) are significant at the 0.1% level, provide an interesting story of the impact of precipitation on irrigation rates. With higher precipitation rates, one might assume that less irrigation would be required to optimize corn yields. But the actual impact is somewhat less clear. Higher values of $PC1$ and $PC2$ result from higher levels of precipitation throughout the primary eastern corn growing season months between May and August, as indicated by Table 5. Although given the square and interaction terms, the magnitude of the change will depend on the temperaure and the initial precipitation level. Since the first stage coefficients of $PC1$ and $PC2$ are greater than zero, higher precipitation during any month of the growing season leads to higher average irrigation rates through these components. This may indeed be the case in the west as higher precipitation will generally mean that more water is available for irrigation, and higher irrigation rates will result as a consequence.

Higher precipitation has a mixed impact on $PC4$ depending on the month. If precipitation increases in May and August, $PC4$ will increase, but higher precipitation in June and July result in slightly lower values of $PC4$. Since the coefficient of $PC4$ is negative, increased precipitation in May and August lead to decreased average yearly irrigation rates through $PC4$, but increased precipitation in June and July increase average yearly irrigation rates. Similar results obtain by including additional months in the principal component analysis or by excluding the temperature and precipitation interaction terms.

1.5.3 Robustness Checks

To check the validity of the results under different sets of assumptions, I perform three robustness checks using the panel GLS random effects estimator with principal components, represented by equation (6). I first check the impact of including counties with lower amounts of corn as a percentage of total harvested cropland to see if the assumption of a

70% threshold is sufficient to consider corn farming as the primary consumer of irrigation water in a county. Second, I exclude Nebraska from the regressions. Nebraska is the primary corn-growing state using a prior appropriation system, accounting for one half of the observations in such states over the three years considered. By excluding Nebraska, I can determine if there is something particular to the "corn-husker" state that may be biasing my overall results. Finally, I run the regression using only observations in those states with similar length growing seasons. Including the 24 observations from the counties in states where growing seasons are longer and multiple plantings may occur, would challenge the vailidity of controlling for climate using growing months between April and September. In each of these alternative specifications, the estimated impacts of water law on corn yields are of the same order of magnitude between 20 and 30 bushels per acre higher in prior appropriation counties.

To check the validity of the threshold of 70% corn, I look at the impact of narrowing and expanding the sample by varying the minimum amount of corn as a percentage of total harvested cropland in a county used to filter the data set. If corn is the predominant crop in a county, the total irrigation rate should be fairly representative of the rate used to irrigate corn. If corn is not the predominant crop, we could expect three different irrigation patterns: (1) all crops are irrigated at about the same rate, (2) corn is preferentially irrigated, or (3) other crops are preferentially irrigated. The first two cases, if consistent across counties, would likely still result in higher corn yields in prior appropriation states regardless of the minimum corn percentage used. In the third case, however, we would expect as the percentage of corn in a county decreases, the percentage of irrigation water used for other crops relative to corn should increase. In this situation, increasing irrigation rates in a county may not result in higher corn yields. This is precisely what happens as counties with lower percentages of corn are included in the sample.

Figure 3 shows the impact that water law has on yields, changing the minimum threshold at which counties are included in the analysis from forty-five percent corn to ninety percent corn in five percent increments (these panel GLS with random effects regressions do not include farm size as a regressor as this data was not collected for counties with less than 70%

corn). Between 65% and 90% corn, there is a consistently higher yield in prior appropriation counties of 20 to 22 bushels per acre compared to riparian law counties. The difference in yields begins to drop sharply when counties 60% corn and below are included. The primary explanation for this drop is the likelihood that scarce irrigation water in prior appropriation states is being used on other crops and is therefore not being used to improve corn yields. As evidence that this is the likely reason for this trend, irrigation water allocation simulators show that when water is constrained, it is more profitable to use the water for crops other then corn (e.g. Paudel et al. (2005) show that peanuts and cotton should be preferentially irrigated).

I also checked the impact of changing the minimum corn threshold using the two-stage model in equations (4) and (5) in order to see how irrigation rates were affected by water law at various thresholds. The first stage estimates of $\hat{\eta}$ continue to show statistically significantly higher irrigation rates in prior appropriation counties of between 1.3 and 2.5 kGal/d/ac compared to counties in riparian law states. The second stage estimates of $\hat{\beta}$, however, begin to drop sharply when counties with 60% and less corn are included. When the threshold is reduced to 50% corn, increasing county irrigation rates no longer results in higher corn yields. When the sample includes counties with more than 65% corn, the impact of water law remains consistently between 20 and 25 bushels per acre higher yield in prior appropriation counties using the two-stage model, consistent with the results of the random effects estimations. This result also rules out the potential that the higher yields associated with prior appropriation law is an artifact of some regional difference between eastern and western states. If this were the case, then dryland corn farming in the west would be expected to have higher yields than dryland farm in the east. When the percent of corn in a county is lower, the more likely that the corn is not irrigated. If western agriculture was predisposed to higher corn yields, the difference between eastern and western yields should remain as more counties are included, but in fact the opposite occurs. I also verified this by substituting various monthly rainfall totals as proxies for water law, and either incorporating the water law indicator into the principal component in place of that monthly rainfall total or omitting the water law indicator completely. Since rainfall is highly correlated with the region of the country, if it was a regional effect leading to higher yields, the regression results should not be substantially different by making these changes. These changes did, however, alter the results significantly. I can therefore conclude that I am not simply picking up a regional difference which also happens to be highly correlated to water law.

Table 8 presents comparisons of the base random effects regression, the regression excluding counties in Nebraska, and regression excluding counties in those states without similar growing seasons. The overall increase in yield in prior appropriation counties when the 60 observations from Nebraska counties are excluded, is 24 bushels per acre, versus the increase of 28 in the base regression. As a final robustness check I eliminate the 24 observations from the counties in Arizona, southern California, New Mexico, and Texas. Estimating the model without these observations indicates that prior appropriation counties will have corn yields that are roughly 30 bushels per acre higher than those in riparian law states, similar to the 28 bushels per acre increase estimated in the base regression.

1.6 Conclusions and Extensions

This paper quantifies the impact on corn yields of allowing non-riparians access to surface water for irrigation in states with a riparian common law system of water rights. I compile a county level data set which includes average corn yield, irrigation rate, farm size, and monthly climate data. Using a two-stage model combining climate variables in a principal component analysis, I find a robust and significant relationship between access to surface water and corn yields for counties that are primarily engaged in corn farming. Estimation under various assumptions indicates that counties in states with prior appropriation water law, where use of surface waters for irrigation does not depend on ownership of adjacent land, enjoy approximately 20 to 30 bushels per acre higher average yields than their riparian counterparts in the eastern U.S. Given corn prices in the range of \$4 to \$6 per bushel, an individual farmer in a riparian law state could increase revenues from corn sales on the order of \$80 to \$180 per acre with increased access to surface water irrigation. While this analysis does not encompass the broader implications of the general equilibrium and welfare effects that are likely to occur, it points to the disadvantages that eastern farmers face relative to those in the west, despite a wetter climate seemingly better suited to crop production.

Should eastern rainfall and temperature norms begin to shift toward lower precipitation and higher average temperatures as projected by climate modeling, increased access to irrigation water may become a necessity to maintain yields and therefore the profitability of the primarily rainfed agriculture of the eastern U.S. Further, as water availability in the western U.S. becomes ever more of an issue given the competing and expanding demands on water resources, agriculture may begin to shift to more water rich areas of the country. Eastern states which adopt a more flexible approach to water management and use will be at a competitive agricultural advantage relative to those that do not in taking advantage of this shift. While residential and commercial uses of water also continue to grow in the east, and fights over water such as that between Alabama, Georgia, and Florida are likely to expand, appropriate legal frameworks will also help direct water to higher value uses.

Despite the apparent benefits of expanding water use rights to non-riparians and allowing water trading in the eastern U.S., such changes would be met with skepticism and even some open hostility. While non-riparian farmers would benefit directly, riparian farmers would likely oppose this change as they currently have a distinct advantage over their neighbors. Environmental groups in the eastern U.S. would also be opposed to any further consumptive use of water resources. Alabama, for example, is home to more endangered aquatic species than any other state. Changing current law to allow expanded productive use of the state's waters would face serious legislative and judicial hurdles. The failure of legislation to allow water permit trading in Georgia during the 2003-2004 legislative session provides just one example of the challenges such a change would entail. McNider et al. (2005) propose an interesting approach to allowing expanded uses of water for agriculture while still meeting the legitimate environmental questions of doing so. They advocate on-farm storage systems wherein diversions are allowed to non-riparians during the wetter winter months when eastern water resources are more than adequate to supply expanded usage.

In contrast to environmental groups in the east, western environmental groups may welcome a change in eastern laws that could result in agriculture shifting to the wetter east. The recent trend in the western U.S. is toward more "in-stream" water uses, for both recreation and environmental purposes. If more fertile eastern acres in a wetter climate can be used for agriculture by allowing access to surface water, some of the pressure on western resources could be relieved. Further, reducing western water uses would diminish water storage requirements and could allow the removal of some western dams. This is a high priority for many western environmental groups.

As a future extension of this analysis, it would be of value to incorporate additional climatic data. First, incorporating measurements of humidity and wind speed, provided these are available, could improve regression results. Lower humidity and higher wind speeds have been shown to decrease soil moisture and therefore increase the water requirement to optimize yields. The effects of humidity, however, are likely to already be captured to some extent in the measures of temperature and precipitation used presently. Another climatic feature of interest would be the application of a modified "growing degree days" measure as in Schlenker and Roberts (2006) (see note 5) rather than relying on a quadratic function of average temperature. Using average monthly temperatures would tend to miss relatively short periods within a month or even within a day where temperature extremes may damage plant growth and reduce overall yields. Schlenker and Roberts' measure accounts for this by generating a composite value over the entire growing season related to the amount of time crops spend at discretized temperature levels.

Appendix A: Figures

Figure 1 – *U.S. Water Laws by State*

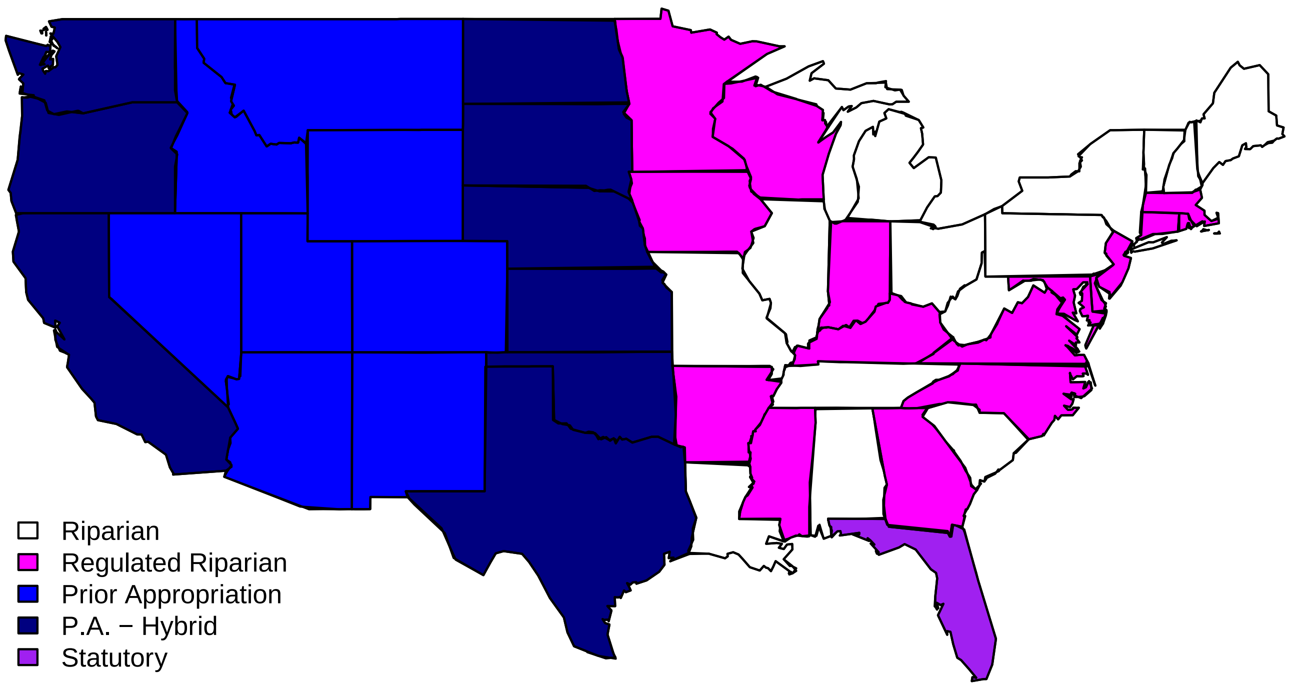
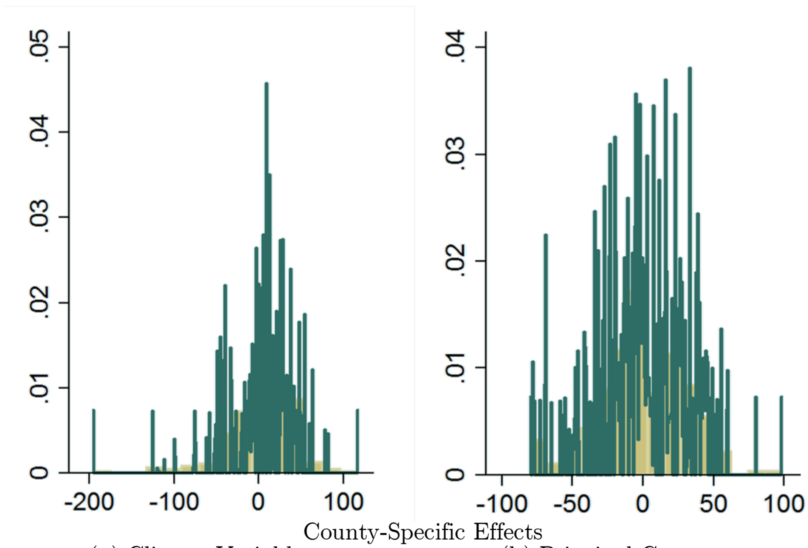
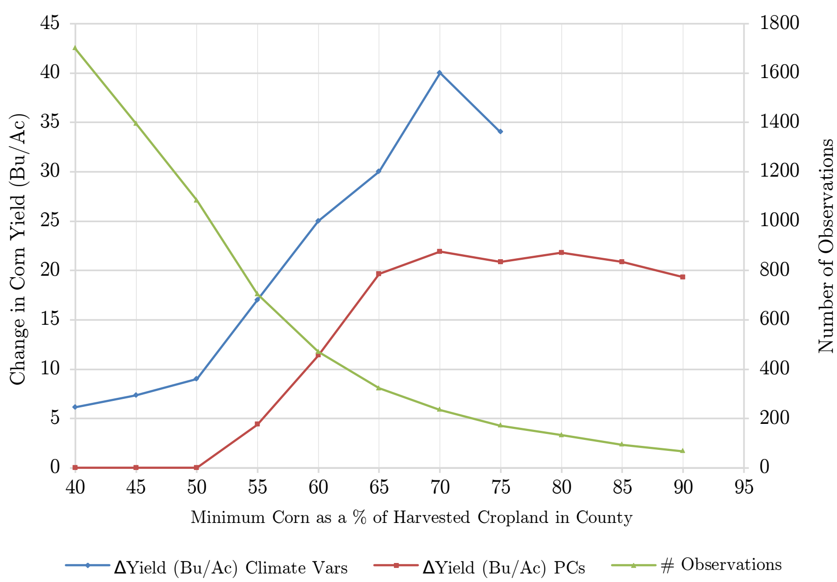


Figure 2 – *Estimated density of the estimated county-specific effects, $\{\hat{\alpha}_i\}_{i=1}^{234}$, from fixed effects regressions using (a) climate variables, and (b) principal components.*



(a) Climate Variables (b) Principal Components

Figure 3 – *Random Effects Estimates of the Effect of Law on Corn Yields by Varying Percent Corn*



Appendix B: Summary Statistics and Regression Results

Table 1 – *Summary Statistics for Counties with greater than 70% Corn (1995, 2000, 2005)*

Variable	Mean	Std. Dev.	Min.	Max.
Corn Yield (Bu/Ac)	118	32.9	40	215
Avg Irr/Ac (kGal/d/ac)	1.58	2.50	0	26.5
Avg Farm Size (100AC)	7.92	16.8	0.64	211
Apr Precip. (in.)	3.12	2.39	0.06	17.1
May Precip. (in.)	5.07	3.87	0.01	23.3
Jun Precip. (in.)	4.50	6.71	0.02	84.9
Jul Precip. (in.)	4.10	7.30	0	65.5
Aug Precip. (in.)	4.88	11.75	0	156
Sep Precip. (in.)	4.00	7.91	0	59.4
Avg Apr Temp. ($^{\circ}F$)	50.1	7.4	35.0	73.6
Avg May Temp. ($^{\circ}F$)	59.2	6.9	45.6	81.7
Avg Jun Temp. ($^{\circ}F$)	69.6	4.6	56.4	86.9
Avg Jul Temp. ($^{\circ}F$)	74.5	4.1	61.5	87.6
Avg Aug Temp. ($^{\circ}F$)	74.7	4.5	61.1	87.8
Avg Sep Temp. ($^{\circ}F$)	65.0	5.8	53.9	86.5
Observations		234		
Counties		170		
States		18		

Table 2 – *Summary Statistics by type of Water Law for Counties with greater than 70% Corn (1995, 2000, 2005)*

Variable	Overall Mean	Prior Approp Mean	Riparian Mean
Corn Yield (Bu/Ac)	118	128	110
Avg Irr/Ac (kGal/d/ac)	1.58	2.37	0.78
Avg Farm Size (100AC)	7.92	12.7	1.88
Apr Precip. (in.)	3.12	2.22	3.72
May Precip. (in.)	5.07	4.43	5.45
Jun Precip. (in.)	4.50	2.61	5.44
Jul Precip. (in.)	4.10	1.68	5.42
Aug Precip. (in.)	4.88	1.84	5.88
Sep Precip. (in.)	4.00	1.80	4.99
Avg Apr Temp. ($^{\circ}F$)	50.1	49.6	50.3
Avg May Temp. ($^{\circ}F$)	59.2	58.8	59.6
Avg Jun Temp. ($^{\circ}F$)	69.6	68.8	70.5
Avg Jul Temp. ($^{\circ}F$)	74.5	75.6	73.8
Avg Aug Temp. ($^{\circ}F$)	74.7	75.7	74.2
Avg Sep Temp. ($^{\circ}F$)	65.0	65.7	64.4
Observations	234	105	129
Counties	170	71	99
States	18	8	10

Table 3 – *Impact of Water Law on Corn Yields: 2SLS, Random Effects, and Pooled OLS Regression Results Using Climate Variables*

	2SLS 1^{st} **Stage**	
Water Law, $\hat{\eta}$	1.94	$(0.57)^{***}$
Avg Farm Size (100AC)	−0.020	$(0.0088)^{*}$
Year = 2000	−1.25	$(0.56)^{*}$
Year = 2005	−0.33	(0.67)
Apr Precip. (in.)	0.34	(0.78)
May Precip. (in.)	−1.74	$(0.73)^{*}$
Jun Precip. (in.)	−0.75	(1.05)
Jul Precip. (in.)	3.23	$(1.42)^{*}$
Aug Precip. (in.)	0.016	(0.85)
Sep Precip. (in.)	−1.10	(0.93)
Avg Apr Temp. ($^{\circ}F$)	0.29	(0.54)
Avg May Temp. ($^{\circ}F$)	−0.31	(0.65)
Avg Jun Temp. ($^{\circ}F$)	−2.92	$(1.28)^{*}$
Avg Jul Temp. ($^{\circ}F$)	2.66	(1.97)
Avg Aug Temp. ($^{\circ}F$)	0.047	(1.52)
Avg Sep Temp. ($^{\circ}F$)	1.06	(1.04)
Apr P^2	0.077	$(0.018)^{***}$
May P^2	0.017	$(0.009)^{+}$
Jun P^2	0.007	$(0.003)^{*}$
Jul P^2	0.003	(0.002)
Aug P^2	−0.002	$(0.001)^{+}$
Sep P^2	−0.001	(0.003)
Apr T^2	−0.002	(0.005)
May T^2	0.002	(0.005)
Jun T^2	0.020	$(0.009)^{*}$
Jul T^2	−0.018	(0.013)
Aug T^2	0.0005	(0.010)
Sep T^2	−0.009	(0.008)
Constant	−26.3	(50.1)
R^2	0.52	
Instrument F	11.5^{***}	
Instrument partial R^2	0.055	

Impact of Water Law on Corn Yields: 2SLS, Random Effects, and Pooled OLS Regression Results Using Climate Variables, cont'd

	2SLS 2^{nd} **Stage**		**GLS, RE**		**OLS**	
Water Law			43.3	(7.39)***	46.0	(6.98)***
Avg Irr/Ac(kGal/d/ac), $\hat{\beta}$	23.8	(5.70)***				
Avg Farm Size (100AC)	0.12	(0.20)	−0.36	(0.05)***	−0.36	(0.06)***
Year = 2000	37.3	(11.7)**	11.3	(6.64)+	7.52	(6.81)
Year = 2005	15.1	(12.9)	13.1	(7.58)+	7.27	(9.01)
Apr Precip. (in.)	8.17	(18.7)	12.2	(7.85)	16.2	(8.67)+
May Precip. (in.)	46.3	(15.8)**	9.61	(6.80)	4.87	(8.24)
Jun Precip. (in.)	10.5	(28.7)	−13.0	(12.5)	−7.43	(12.7)
Jul Precip. (in.)	−56.0	(49.1)	14.7	(13.5)	20.6	(14.5)
Aug Precip. (in.)	−11.9	(22.5)	−13.4	(8.31)	−11.6	(8.80)
Sep Precip. (in.)	32.7	(22.8)	10.7	(8.30)	6.47	(8.64)
Avg Apr Temp. ($°F$)	−8.09	(9.48)	1.26	(6.19)	−1.16	(6.23)
Avg May Temp. ($°F$)	24.2	(13.7)+	11.2	(7.03)	16.8	(7.70)*
Avg Jun Temp. ($°F$)	3.10	(14.3)	−13.5	(16.3)	55.8	(28.9)*
Avg Jul Temp. ($°F$)	−23.5	(47.2)	20.1	(23.5)	39.7	(27.3)
Avg Aug Temp. ($°F$)	−8.85	(32.3)	−4.70	(17.8)	−7.74	(20.0)
Avg Sep Temp. ($°F$)	−12.2	(22.8)	10.4	(10.8)	13.02	(14.0)
Apr P^2	−1.97	(0.96)*	0.024	(0.19)	−0.13	(0.18)
May P^2	−0.40	(0.20)*	−0.041	(0.081)	0.015	(0.090)
Jun P^2	−0.15	(0.11)	−0.003	(0.030)	0.002	(0.030)
Jul P^2	−0.088	(0.11)	−0.009	(0.018)	−0.010	(0.020)
Aug P^2	0.052	(0.038)	0.005	(0.009)	0.009	(0.010)
Sep P^2	0.020	(0.098)	−0.002	(0.030)	−0.007	(0.030)
Apr T^2	0.082	(0.093)	0.0098	(0.063)	0.035	(0.062)
May T^2	−0.19	(0.11)+	−0.097	(0.057)	−0.14	(0.061)*
Jun T^2	−0.35	(0.19)+	0.0021	(0.11)	0.13	(0.12)
Jul T^2	0.16	(0.31)	−0.15	(0.16)	−0.27	(0.18)
Aug T^2	0.007	(0.22)	0.006	(0.12)	0.020	(0.13)
Sep T^2	0.10	(0.17)	−0.08	(0.085)	−0.10	(0.11)
Constant	−875	(1100)	−1240	(560)*	−1500	(548)**
Observations	234		234		234	
R^2	.		0.66		0.66	
F	6.28***				35.2***	
Wald χ^2			1060***			
Breusch-Pagan LM			18.5***			
Hausman Test: RE vs FE			28.3			
Hausman Test: Irr.Endog.	104					
Poolability F Test					0.04	

Robust standard errors clustered at the county level, in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

RE and OLS estimates are from equation (3) while 2SLS estimates are from equations (1) and (2). Interaction terms are not reported. The bootstrap Hausman test of irrigation rate endogeneity using robust standard errors with 400 replications is reported.

Table 4 – *Climate Data Correlation Matrix*

	Apr P	May P	Jun P	Jul P	Aug P	Sep P	Apr T	May T	Jun T	Jul T	Aug T
May Precip (in.)	0.80										
Jun Precip (in.)	0.75	0.79									
Jul Precip (in.)	0.81	0.75	0.79								
Aug Precip (in.)	0.79	0.68	0.73	0.78							
Sep Precip (in.)	0.75	0.73	0.77	0.79	0.70						
Avg Apr Temp ($°F$)	−0.13	−0.17	−0.09	−0.12	−0.10	−0.11					
Avg May Temp ($°F$)	−0.12	−0.13	−0.08	−0.14	−0.11	−0.11	0.87				
Avg Jun Temp ($°F$)	−0.06	−0.14	−0.09	−0.05	−0.03	−0.06	0.83	0.77			
Avg Jul Temp ($°F$)	−0.20	−0.24	−0.20	−0.19	−0.16	−0.22	0.84	0.77	0.84		
Avg Aug Temp ($°F$)	−0.18	−0.19	−0.20	−0.20	−0.17	−0.25	0.73	0.74	0.76	0.90	
Avg Sep Temp ($°F$)	−0.13	−0.20	−0.11	−0.11	−0.11	−0.12	0.92	0.81	0.86	0.87	0.76

Table 5 – *Variable Weights from Principal Component Analysis*

	Comp 1	**Comp 2**	**Comp 3**	**Comp 4**
Apr Precip (in.)	0.231	0.100	−0.045	0.204
May Precip (in.)	0.226	0.081	0.294	0.251
Jun Precip (in.)	0.228	0.103	0.133	−0.104
Jul Precip (in.)	0.232	0.103	−0.089	−0.093
Aug Precip (in.)	0.219	0.104	−0.346	0.109
Sep Precip (in.)	0.225	0.094	0.075	−0.315
Avg Apr Temp ($^{\circ}F$)	−0.118	0.271	0.039	−0.075
Avg May Temp ($^{\circ}F$)	−0.113	0.258	0.102	−0.014
Avg Jun Temp ($^{\circ}F$)	−0.105	0.270	−0.059	−0.036
Avg Jul Temp ($^{\circ}F$)	−0.141	0.256	−0.035	0.054
Avg Aug Temp ($^{\circ}F$)	−0.135	0.236	−0.002	0.184
Avg Sep Temp ($^{\circ}F$)	−0.121	0.271	−0.001	−0.082
Apr P^2	0.207	0.095	−0.056	0.243
May P^2	0.185	0.077	0.417	0.324
Jun P^2	0.183	0.090	0.174	−0.151
Jul P^2	0.204	0.104	−0.119	−0.145
Aug P^2	0.169	0.094	−0.489	0.143
Sep P^2	0.189	0.090	0.086	−0.426
Avg Apr T^2	−0.118	0.273	0.034	−0.071
Avg May T^2	−0.113	0.259	0.098	−0.022
Avg Jun T^2	−0.106	0.272	−0.056	−0.040
Avg Jul T^2	−0.141	0.258	−0.030	0.052
Avg Aug T^2	−0.135	0.238	0.003	0.180
Avg Sep T^2	−0.121	0.271	−0.002	−0.079
Apr $P \times T$	0.222	0.121	−0.054	0.196
May $P \times T$	0.220	0.093	0.312	0.264
Jun $P \times T$	0.225	0.112	0.118	−0.111
Jul $P \times T$	0.227	0.111	−0.104	−0.094
Aug $P \times T$	0.213	0.112	−0.363	0.119
Sep $P \times T$	0.219	0.105	0.063	−0.336
Eigenvalues, λ	13.7	9.48	1.35	1.15
Proportion Variation, λ/n	0.46	0.32	0.05	0.04

Table 6 – *Impact of Water Law on Corn Yields: 2SLS, Random Effects, and Pooled OLS Regression Results Using Principal Components*

	2SLS 1^{st} Stage	
Water Law, $\hat{\eta}$	2.51	$(0.33)^{***}$
Avg Farm Size (100AC)	-0.011	(0.0093)
Year = 2000	-0.68	(0.44)
Year = 2005	-0.47	(0.38)
Component 1	0.68	$(0.06)^{***}$
Component 2	0.13	$(0.03)^{***}$
Component 3	0.0003	(0.33)
Component 4	-2.93	$(0.71)^{***}$
Constant	1.37	$(0.30)^{***}$
R^2	0.26	
Instrument F	58.5^{***}	
Instrument partial R^2	0.21	

	2SLS 2^{nd} Stage		GLS, RE		OLS	
Water Law			28.3	$(4.54)^{***}$	26.8	$(4.61)^{***}$
Avg Irr/Ac (kGal/d/ac), $\hat{\beta}$	10.7	$(2.66)^{***}$				
Avg Farm Size (100AC)	−0.37	$(0.061)^{***}$	−0.47	$(0.06)^{***}$	−0.48	$(0.06)^{***}$
Year = 2000	20.8	$(7.14)^{***}$	13.3	$(4.05)^{***}$	13.5	$(4.90)^{**}$
Year = 2005	33.0	$(6.69)^{***}$	25.9	$(3.80)^{***}$	27.9	$(4.41)^{***}$
Component 1	−11.7	$(1.79)^{***}$	−3.28	$(0.60)^{***}$	−4.43	$(0.54)^{***}$
Component 2	−7.62	$(1.06)^{***}$	−5.84	$(0.47)^{***}$	−6.18	$(0.48)^{***}$
Component 3	−14.1	$(6.56)^{***}$	−13.8	$(3.44)^{***}$	−14.1	$(3.59)^{***}$
Component 4	11.4	(18.5)	−26.2	$(6.86)^{***}$	−19.9	$(6.97)^{***}$
Constant	76.0	$(8.99)^{***}$	92.3	$(3.06)^{***}$	90.7	$(2.81)^{***}$
Observations	234		234		234	
R^2	.		0.45		0.45	
F	13.0^{***}				27.9^{***}	
Wald χ^2			313^{***}			
Breusch-Pagan LM			30.6^{***}			
Hausman Test of RE vs FE			12.2			
Hausman Test of Irr. Endog.	655^{***}					
Poolability F Test					0.83	

Robust standard errors clustered at the county level, in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

RE and OLS estimates are from equation (6) while 2SLS estimates are from equations (4) and (5).
The bootstrap Hausman test of irrigation rate endogeneity using robust standard errors with 400 replications is reported.

Table 7 – *Impact of Water Law on Corn Yields: Random Effects Regressions of Single Stage Model With Climate Variables(CVs) and Principal Components(PCs)*

	GLS, RE (CVs)		**GLS, RE (PCs)**	
Water Law	43.3	(6.46)***	28.3	(4.39)***
Avg Farm Size (100AC)	−0.36	(0.10)***	−0.47	(0.11)***
Year = 2000	11.3	(5.61)*	13.3	(3.95)***
Year = 2005	13.1	(6.94)+	25.9	(3.49)***
Component 1			−3.28	(0.60)***
Component 2			−5.84	(0.47)***
Component 3			−13.8	(3.44)***
Component 4			−26.2	(6.86)***
Apr Precip. (in.)	12.2	(8.58)		
May Precip. (in.)	9.61	(7.91)		
Jun Precip. (in.)	−13.0	(10.8)		
Jul Precip. (in.)	14.7	(14.8)		
Aug Precip. (in.)	−13.4	(9.42)		
Sep Precip. (in.)	10.7	(10.2)		
Avg Apr Temp. ($°F$)	1.26	(6.06)		
Avg May Temp. ($°F$)	11.2	(6.91)		
Avg Jun Temp. ($°F$)	3.10	(13.8)		
Avg Jul Temp. ($°F$)	20.1	(18.9)		
Avg Aug Temp. ($°F$)	−4.70	(14.8)		
Avg Sep Temp. ($°F$)	10.4	(11.2)		
Apr P^2	0.024	(0.20)		
May P^2	−0.041	(0.10)		
Jun P^2	−0.003	(0.035)		
Jul P^2	−0.009	(0.024)		
Aug P^2	0.005	(0.011)		
Sep P^2	−0.002	(0.031)		
Apr T^2	0.010	(0.06)		
May T^2	−0.097	(0.056)+		
Jun T^2	0.002	(0.100)		
Jul T^2	−0.15	(0.130)		
Aug T^2	0.006	(0.100)		
Sep T^2	−0.081	(0.085)		
Constant	−1240	(560)*	92.3	(3.06)***
Observations	234		234	
R^2	0.66		0.45	
Wald χ^2	1060***		313***	
Breusch-Pagan LM	18.5***		30.6***	
Hausman Test of RE vs FE	28.3		12.2	

Robust standard errors clustered at the county level, in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

RE estimates using climate variables are from equation (3), and RE estimates using principal components are from equation (6). Interaction terms are not reported.

Table 8 – *Base RE Results (1) Compared to RE Regressions without Nebraska (2) and without California and Southwestern States (3)*

	(1) Base RE Model		**(2)** RE Without Nebraska		**(3)** RE Without CA & SW	
Water Law	28.3	(4.54)***	23.7	(6.89)***	30.7	(4.01)***
Avg Farm Size (100AC)	−0.47	(0.06)***	−0.45	(0.06)***	−0.41	(0.05)***
Year = 2000	13.2	(4.05)***	14.3	(5.10)***	13.1	(3.82)***
Year = 2005	25.9	(3.80)***	22.2	(4.11)***	29.0	(3.80)***
Component 1	−3.28	(0.60)***	−4.44	(0.66)***	−2.26	(0.52)***
Component 2	−5.84	(0.47)***	−6.29	(0.49)***	−2.26	(0.49)***
Component 3	−13.8	(3.44)***	−16.9	(3.73)***	−8.99	(2.94)***
Component 4	−26.2	(6.86)***	−31.2	(7.33)***	−24.0	(6.19)***
Constant	92.3	(3.06)***	92.3	(3.15)***	95.8	(3.06)***
Observations	234		174		215	
Wald χ^2	312***		209***		387***	

Robust standard errors clustered at the county level, in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2 The Effect of Production Royalties on a Non-Renewable Resource: Oil Production from Marginal Wells

Abstract

Two questions addressed in this study consider to what extent oil producers respond to changes in price and whether higher royalties on oil producers result in a reduction in the life of a producing lease. This paper uses a unique federal lease-level data set to estimate the response by oil producers to changing monthly prices between 1990 and 2011. In addition to the normal changes in price on the world market, variation is also achieved by using the actual price that each producer received for oil sales each month, rather than an average annual benchmark price of oil of the type used in most studies. Further, marginal producers, which make up approximately 28% of federal leases, were afforded royalty relief between 1992 and 2006, imparting additional price variation. I find that production from these marginal volume leases is relatively more elastic than other federal leases, regardless of the specification. Marginal leases with regularly reported sales have a zero or small positive supply elasticity depending on which price is used. The broader pool of leases not classified as marginal producers have a small, negative, generally insignificant supply elasticity. Further, leases that participated in the royalty reduction program have a roughly 15% lower probability of being shut-in than those leases that were not eligible and were required to pay the full royalty amount on production.

2.1 Introduction

Recent policy proposals by the federal executive branch, supported by polls of the American public at large,[13] call for the elimination of various incentives in the tax code aimed at reducing oil exploration and production costs. The cost of these incentives is estimated to be in the neighborhood of $4 billion in lost government revenue annually.[14] On the other hand, in an attempt to generate increased revenues from oil production, many states have increased the production royalties for new leases on state lands or have increased

[13]See for example, Question 26 of the NBC News/Wall Street Journal Survey dated February 2011 available at http://online.wsj.com/public/resources/documents/wsj-nbcpoll03022011.pdf

[14]See for example, the President's Fiscal Year 2013 Budget of the U.S. Government, p. 44, available at http://www.whitehouse.gov/sites/default/files/omb/budget/fy2013/assets/budget.pdf.

the severance taxes on oil production in the state. To assess the potential impact of such policies, an understanding of how oil producers respond to changes in prices, including taxes, is required. The present study examines this issue using micro-level data.

I investigate to what extent oil producers respond to changes in price, and whether higher royalties on oil producers, an *ad valorem* tax, result in a reduction in the life of a producing lease. To do so, I use a unique monthly data set of federal leases of wells producing marginal volumes of oil between January 1990 and May 2011, as well as data for all other federal leases beginning in January 1996, to estimate the price elasticity of oil supply. Two key features of the data make it ideal for this task. First, the data set contains the actual monthly sales price of oil from each lease and the actual royalty rate under which the lease was operating each month. This provides variation within and across leases over time rather than requiring generalizations based on benchmark prices and aggregate annual U.S. production. Second, the impact of price on marginal volume, or "stripper," wells can be inferred from the data on leases which qualified for the Bureau of Land Management's (BLM) Stripper Oil Well Property Royalty Rate Reduction Program between October 1992 and February 2006. To qualify for the program, the lease was required to produce less than fifteen barrels per day per well on average, and depending on how much below the threshold it produced in a qualifying period, the royalty rate was reduced accordingly. Overall, oil stripper wells on both public and private lands accounted for more than 16% of U.S. oil production in 2009 (U.S. Energy Information Administration, 2011), and any program designed to influence domestic oil production could have a potentially large impact through its effect on these wells.

The study contributes to two important areas of resource economics. The first contribution is to provide some clarity to a set of conflicting prior estimates of the price elasticity of oil production. Virtually all existing studies rely on aggregate annual U.S. production and average annual "oil" prices.[15] Crude oil, however, is not a homogeneous resource and different types of oil, defined generally by their so-called API gravity and sulfur content, command vastly different prices. Further, producers face widely varying production costs

[15] The only exception I am aware of is Rao (2010), which is unpublished (see Section 2.3 for a discussion).

depending on factors such as the geology of the formation, API gravity of the crude, how long it has been producing, and the transportation costs from the well to the market. Perhaps as a result of neglecting such differences, production studies relying on aggregate U.S. data have found a wide range of supply elasticities, including negative short run elasticities (for a review of these studies, see Section 2.3). For the purpose of making policy decisions, these conflicting estimates are unhelpful. The present study, on the other hand, uses lease-level data and thus allows for differences in product characteristics and production costs to be picked up in fixed effects regressions, resulting in more accurate estimations of the supply elasticity.

The second contribution is to perform an empirical analysis of the impact of taxation on the supply of exhaustible resources, in the context of U.S. stripper wells. With constant marginal extraction costs in a perfectly competitive market without uncertainty, the Hotelling rule states that the asset price of a non-renewable resource will equal the scarcity rent and grow at the rate of interest. Taking depletion of the resource into account, the shadow price increases at a rate slower than the rate of interest "because extraction today leads to higher costs tomorrow, and the owner internalizes this externality" (Slade and Thille 2009, p. 245). For most exhaustible resources, however, the growth of prices predicted by the Hotelling rule has not been realized, even accounting for the impact of depletion (see for example Gaudet 2007, and Figures 4 and 5). One possible explanation is that resource taxation impacts prices and extraction decisions (Gamponia and Mendelsohn 1985; Yucel 1988; Sweeney 1993; Gaudet 2007; Slade and Thille 2009). While theory tells us how the extraction and price path should respond to taxation (Sweeney 1993; Dasgupta and Heal 1979), much remains to be done in studying how particular industries actually respond. In this paper I focus on the operation of oil leases, and in particular marginal volume well leases.

Oil field operation differs from other exhaustible resource operations in that it is generally governed more by geology and the mechanics of the extraction process than by prices. Once production commences from a lease, it will typically produce a maximum amount that tapers off as the pressure in the reservoir declines (e.g., Kunce et al. 2003, p. 5). If, at this time, the market price less royalties and taxes is above some threshold and forecasted to remain so, the operator will employ additional and more costly recovery techniques to extend the producing life of the reservoir. So long as the effective price remains above the higher marginal cost of maintaining reservoir pressure, the lease will continue to produce, but at a decreasing rate. However, if the price falls too much, or costs to maintain the same rate of production rise too much, the wells will be plugged and the field abandoned. Plugged wells will generally not be brought back into service given the costs associated with unplugging, reestablishing flow, and maintaining reservoir pressure. Therefore, the decision to plug wells on a lease and stop production is not taken lightly and may lead to production responses unlike those seen for mined resources where production rates can be adjusted readily. Therefore, the only price responses of oil producers from currently producing leases in the short term is the decision whether or not to shut wells to reduce production, or to abandon an oil field altogether.

The BLM's Stripper Oil Well Property Royalty Rate Reduction Program provides an excellent vehicle with which to test producer responses to price changes for a subset of marginal leases that are theoretically the most responsive within the broader pool of leases. Five key results are of note. First, production from leases that qualified as federal stripper oil well leases is relatively more elastic than production from non-stripper well leases. Second, stripper well leases with regularly reported sales (generally those with higher production) have an estimated supply elasticity between 0 and 0.07. Non-stripper well leases with regularly reported sales are relatively less elastic, with generally insignificant elasticities estimated between 0 and -0.1.[16] Third, the most marginal of the qualified leases have a supply elasticity estimated to be between 0.09 and 0.56, while the most marginal of the non-stripper well leases have an estimated supply elasticity of 0. Long run elasticity estimates are also consistent with these results, with two important differences: (1) the most marginal of non-stripper well leases have a small positive long-run elasticity of 0.06, and (2) the most marginal of the stripper well leases have a long-run elasticity of 0.62. Finally, a lease survival analysis indicates that leases that participated in the royalty re-

[16]These negative estimates are not robust to changes in the level of aggregation or to which price is used (sale price or lagged sale price), and do not imply a negative price elasticity for domestic oil supply.

duction program have an approximately 15% lower probability of being shut-in than those leases that did not participate. Given the limited time-frame over which data is available for non-stripper well leases, which results in the lack of an appropriate counterfactual, the last result should be viewed with caution.

These results have important policy implications related to oil taxation. First, by reducing production royalties for the most marginal of leases during periods of lower prices (increasing the effective price that the producers receive for the oil), producers can be expected to respond by investing to extend the producing life of reservoirs. This result is predicted by theory, and the empirical analysis provides confirmation using an ideally suited data set. Second, higher production royalties do not appear to deter production for the bulk of federal leases. This suggests that additional revenue could be raised in the short term by increasing royalties on federal lands from the current $12\frac{1}{2}\%$, for example to a level similar to that for production on state lands (generally $16\frac{2}{3}\%$). Given that my estimates for long-run elasticities for currently producing non-stripper well leases are also approximately zero, the analysis suggests that producers of these wells are basing operation decisions on something other than price. It is likely that once production commences, production rates are based on geological features of the formation as well as the management of the reservoir for maximum lifetime recovery. Only towards the end of a lease's producing life does price appear to affect production decisions. Thus, exploration and development decisions (i.e., decisions in the "very long-run") appear then to be the main production response to changing prices. Finally, it is likely that the BLM's royalty reduction program for stripper well leases was too inclusive and a lower production threshold could have been used for lease eligibility in the program.

This paper is organized as follows. Section 2.2 provides a discussion of the BLM's Stripper Oil Well Property Royalty Rate Reduction Program. Section 2.3 discusses the previous studies that attempt to estimate oil supply elasticities, focusing primarily on those studies using U.S. data. Sections 2.4 and 2.5 describe the estimation strategy and sources of data used in this analysis, respectively. Section 2.6 discusses empirical results followed by conclusions and extensions in Section 2.7.

2.2 Stripper Well Royalty Reduction Program

The Stripper Oil Well Property Royalty Rate Reduction Program[17] was designed to:

> Provide an economic incentive for operators to restart production of marginal or uneconomic oil wells and to increase production on Federal onshore leases by drilling new wells, by reworking existing wells, and/or by implementing enhanced or secondary oil recovery projects (LaRouche 2001, p. 5).

Royalty rates for leases on federal properties are generally defined at a fixed $12\frac{1}{2}\%$ interest. The lessee must pay the BLM $12\frac{1}{2}\%$ of the value of the oil sold, or transfer to the BLM $12\frac{1}{2}\%$ of the lease production on a monthly basis. Between October 1992 and February 2006, royalty rates were reduced based on the schedule of rates in Appendix B.

Importantly, rates for the entire period were determined by the average per well production in the twelve-month qualifying period between August 1990 and July 1991, a period of time prior to enactment of the program. This prevented producers from lowering their production intentionally for a year in order to shift production to later years at a lower royalty. If a lease was not producing during that period, the qualifying period became the twelve months immediately preceding the shut-in of the wells on the lease. If average per well production of the lease continued to decline, participating leases could reapply to receive lower rates. However, if production subsequently increased, the royalty reverted to the initial qualifying rate. As the program was designed to stimulate production from marginal leases, even if the average production in years subsequent to the qualifying period exceeded the threshold on which the initial qualifying rate was based, producers were still only required to pay the initial rate. Indeed, LaRouche (2001) surveyed the 100 largest benefiting leases in the program and found that 43% exceeded the fifteen BPD per well annual average in 2001.

The program was discontinued by the BLM on February 1, 2006 with six months prior notice. The termination resulted from a provision in the regulation allowing BLM to discontinue the incentive if the price of the benchmark West Texas Intermediate (WTI) crude

[17] 43 CFR 3103.4–2

exceeded \$28 per barrel for six consecutive months (adjusted for inflation with 1991 as the base year). This threshold was exceeded for the six consecutive months ending July 2005 when notice of termination was given by BLM (the inflation adjusted threshold was approximately \$36 per barrel). Overall, approximately 4,700 leases participated in the program, 40% of which were in New Mexico with another 40% in Wyoming. While the number of individual operators participating over the lifetime of the program has not been published, as of October 1999 that number stood at roughly 850. From the beginning of the program until September 1999, approximately half the lifetime of the program, the total amount of royalty relief was more than \$139 million (1999 \$).

2.3 Literature Review

2.3.1 Supply Elasticity of Oil

To analyze the effectiveness of such a program in fulfilling the stated goals of encouraging production on marginal producing properties, an estimate of oil supply elasticity is required. The literature, unfortunately, does not provide much guidance in creating an accurate assessment. Early studies of domestic production yielded supply elasticity estimates that ranged from 0 (MacAvoy 1982) to 2 (Mancke 1970). Studies conducted since the early 1970s vary from supply elasticities of -0.08 (Dahl and Yucel 1991) to 0.26 (Rao 2010). Several authors have found negative short-run supply elasticities (Griffin 1985; Dahl and Yucel 1991; Krichene 2005).

Error Correction and Cointegration: Global Supply Krichene (2005) estimates world supply elasticity using two different econometric techniques. He estimates a system of simultaneous supply and demand equations using an error correction model (ECM) to obtain short-run elasticities and an error correction and cointegration method to obtain long-run elasticities. In the short-run, the ECM indicates that global oil supply is price inelastic (-0.03). The cointegration model indicates that global production response is relatively more elastic in the long-run (0.25). While these insights are important in assessing macroeconomic trends and global markets, they are less instructive for assessing the impact of domestic policy proposals. Further, he was studying production over a very long period from 1918 through 2004. The negative sign is particularly troubling for a supply elasticity, however, the world market can generally not be considered competitive given the control over price that OPEC has exerted through it's supply decisions. Kirchene explains that

> ...the negative sign of the supply price elasticity may derive from the short-run price inelasticity of the demand function. Recognizing the inelastic nature of short-run demand, producers may deliberately refrain from increasing output at the time of a price rise in order to preserve gains in prices ... (2005, p. 10).

However, this is unlikely to be the case for domestic producers of crude oil. Given the competitive pressures from the large number of individual U.S. leaseholders in the present study, a negative elasticity would not be expected.

Residual Supply Curve Approach: Aggregate U.S. Data Mancke (1970) attempts to use inferred estimates of oil industry costs to characterize the shape and level of the industry supply curve for the United States. Production costs in the industry are comprised of the costs of exploration, development, and operation of wells. Using a "sum of components" method would entail adding together the expected values of each of the three cost elements for each barrel of produced oil, although the data requirements for such a study would obviously make it impractical. Instead, Mancke attempts to estimate from available annual production and price data between 1955 and 1968 the percentage of crude oil produced with costs that are near perfectly price-elastic. From this estimate, and an estimate of total expected rents, he determines that the lower bound of supply elasticity must be greater than one. At the time, the U.S. had import quotas restricting the amount of oil that could be obtained from foreign sources at a cost below that which crude oil could be produced domestically. Mancke's analysis was an attempt to estimate the costs of the import quotas to the U.S. in the form of higher oil prices compared to Canada's lower price. Unfortunately, this method only allows broad inferences as to the range of oil supply elasticity and do not apply to current domestic production. Mancke's analysis is based on an expanding level of domestic production when most major producing areas of the country were not in decline, unlike the current reality of domestic production.

Two-Stage Least Squares Joint Supply and Demand Estimation: Aggregate U.S. Data MacAvoy (1982) compares the results from OLS regression of annual oil supply data between 1955 and 1973 to a two-stage least squares model with both demand and supply equations. His supply model involves a regression on price, reserves, and lagged supply. In both cases the coefficient on oil price is insignificant implying an elasticity of zero. As summarized in Dahl and Duggan (1996) he accounts for simultaneity of the supply and demand equations without finding significant differences with a supply model alone. MacAvoy unfortunately does not account for the U.S. price controls or factors such as depletion and technological change.

Ordinary Least Squares Estimation: Aggregate U.S. Data Griffin (1985) investigates the market structures of both OPEC and non-OPEC countries. The basic model for a cartel structure is:

$$lnQ_{it} = \alpha_i + \gamma_i lnP_t + \beta_i lnQ_{-it}^{OO} + \epsilon_{it}$$

where Q_{it} represents annual production in each OPEC country i at time t, P_t is the real price of oil, and Q_{-it}^{OO} is the amount of production in OPEC countries other than country i. In this context, evaluation of different country market structures is achieved by estimating the model individually for each country and then testing of various combinations of values for γ_i and β_i. Using quarterly production and price data from t = 1971 through 1983 he finds that the partial market sharing hypothesis that was anticipated for this group of countries could not be rejected for all 11 countries ($\beta_i > 0$ and $\gamma_i \neq 0$). His analysis of 11 non-OPEC countries proceeds in a similar fashion; however, a competitive market structure model is also considered:

$$lnQ_{it} = \alpha_i + \gamma_i lnP_t + \epsilon_{it}$$

Using available annual data from 1971 through 1982, he finds that a competitive market structure ($\gamma_i > 0$) could only be rejected for the U.S. Again, a significant, negative elasticity is obtained for oil supply. Griffin attributes this to the domestic price controls in effect throughout much of that period.

Jones (1990) conducts an analysis identical to Griffin in order to retest the hypotheses during a period of falling prices with data through the fourth quarter of 1988. Again, the partial market sharing hypothesis is not rejected for the 11 OPEC countries. Importantly, the U.S. market, with the addition of the data after the end of price controls, has a statistically significantly higher supply elasticity (a change of +0.07), indicating a shift towards a competitive market structure.

Dahl and Yucel (1991) approach the question of market structure with an expanded model using a U.S. time series which takes cost data and dynamic decision-making into account. Their model is:

$$lnQ_t = \beta_0 + \beta_p lnP_t + \beta_w lnW + \beta_y lnY + \beta_{qw} lnQ_w + \beta_c lnC + \beta_1 lnI$$

where W is the number of wells drilled, Y is GDP, Q_w is the production in the rest of the world, C is production costs, and I is investment. Their data for non-OPEC countries are annual aggregate production, price, cost, and investment data from $t = 1971$ through 1987. Overall they find that OPEC countries act noncompetitively with some weak evidence of coordination and swing production for some of the countries. While admitting data limitations that yield rather weak results for non-OPEC countries, they find no evidence of dynamic optimization or competitive behavior. They find the supply elasticity for the U.S. to be -0.08 and insignificant, however, a large portion of their U.S. data was also from during the period of price controls. While the approach that Dahl and Yucel (1991) employ is the most comprehensive of the models previously discussed for estimating a domestic oil supply elasticity, the required cost and investment data is not available at the level of the individual producer. This precludes its use in studies which do not employ agregate data sets.

Two additional studies using aggregate data are also worth noting. Using U.S. data from 1958 through 1987, Kandel (1990), as reported in Dahl and Duggan (1996), finds a statistically insignificant coefficient when production level is regressed on price level, indicating an elasticity of 0. Hogan (1989), using U.S. data from 1966 through 1987, regressed the log of production on the log of price averaged over the previous six years, a time trend to account for technological change and reserve depletion, and log of lagged production. He finds a significant short run elasticity of 0.09 and a long-run elasticity of 0.58.

Ordinary Least Squares: California Well Data Finally, Rao (2010) uses a panel of well-level production and field-level oil price data for California monthly production between 1977 and 2008 to assess the likely impact on production decisions of a proposed temporary tax increase. By using well-level panel data, which allows the use of well fixed effects, Rao can account for the production cost heterogeneity in the absence of the required data. With a simple model in levels that controls for well age, month-by-year dummies, and well fixed effects, she finds that the price elasticity of supply for California wells using OLS is between 0.21 and 0.26 depending on the specification.

2.3.2 Taxation of Exhaustible Resources

In addition to the standard taxes on corporate profits and property taxes, two types of taxes have historically been levied on oil producers: excise taxes and *ad valorem* taxes. Profit taxes, or the removal of incentives which would act as if they were higher taxes on the profits of an oil company, would not be expected to impact the short run production decisions of currently operating wells. Instead, such changes would reduce the supply of capital to the industry for the purposes of exploration and development of new fields by reducing the rate of return on capital investment.

Excise taxes would be expected to increase per unit costs and therefore reduce extraction rates. Rao (2010) uses a highly detailed set of California well-level data in an attempt to assess the welfare costs of excise taxes on oil production. Her assessment of the welfare impacts of the tax rely on an assumption of perfect knowledge of the price path by the operator. Slade (1984) attempts to simulate the impacts of various taxes using cost and production data from a copper mine. She finds that royalties decrease copper production rates and ultimately lead to less cumulative extraction. She asserts that the results should generalize to other exhaustible resources. If oil conformed to the results from Slade's copper mine, an increase in royalty rates or severance taxes, which are both *ad valorem* taxes on production,[18] would tend to reduce the amount of domestically produced crude oil. This could be due to either reductions in the rate of extraction, or production lost from those leases that could no longer profitably extract the oil and are shut in with additional recoverable resource left in the ground (i.e. marginal costs become higher than the price of oil less royalty payments and severance taxes). Kunce et al. (2003) cast doubt on the applicability of Slade's conclusion to oil resources. They use aggregate production, reserves, and well data along with information on state taxes to estimate the parameters used in a dynamic simulation model of the impact of severance taxes on Wyoming oil production. They find that long-run production is relatively inelastic with respect to severance tax rate (0.057). This estimate includes the impacts of tax rates on both exploration and production decisions.

2.4 Empirical Analysis

Unfortunately, well-level data as used in Rao (2010) is not the appropriate level of aggregation given how leases or units[19] are operated. Where leases are unitized, investment and production decisions are made at the unit-level and the unit's aggregate production is divided up amongst constituent leaseholders based on the sizes of the leaseholds in relation to the size of the unit. Units or leases can have anywhere from one well to upwards of one thousand wells spread over several thousand acres. Production of an individual well is generally irrelevant. All the wells in a unit or individual lease are operated to maximize

[18]The state severance taxes on oil are generally *ad valorem*, unlike severance taxes on many other minerals which are generally on a per unit extracted basis.

[19]Where individual leases produce from a common reservoir, the leases are often joined into a "unit" with a single operator making production decisions for the unit as a whole. Proceeds from the production are then distributed to the individual leaseholders based on a predetermined division.

the profits of the lessee. Older, shallower wells may be abandoned or left to pump at their capacity if not reworked, while additional wells in the unit or lease may be drilled to take advantage of newer technology. Steps taken to increase the reservoir pressure could, given the geology of the formation, impact production from all wells. Further, if a new well is drilled tapping into the same reservoir, it can impact production of other wells in the vicinity.

Production at the lease-level is collected by the federal government for the calculation of royalties due from the lessees of federal lands. Where leases are operated individually, this is clearly the level of aggregation preferred to accurately reflect producer response. In cases where leases are unitized, the production at the unit-level would be required. Until unit-level data is compiled by the relevant federal and state government agencies, data on production from the constituent leases is the next best alternative. If a unit operator decides to increase production from the unit as a whole by drilling more wells or employing enhanced recovery methods, when the increased production is apportioned to each of the constituent leases each lease will show increased production. As the share of the unit production per lease is fixed, changes in unit production will always be reflected in the changes in individual lease production. The problem with such an approach, however, is that producer response is likely to be more significant than it otherwise would be since several observations will be responding to changes in price with the same percentage response in supply.

2.4.1 Static Panel Analysis

I employ a model similar to that used in Rao (2010), but instead of regression in levels, I consider regressions in logs in order to estimate the elasticity directly. Further, year by month dummies are excluded in the present study. Events influencing all leases each month are limited to either federal tax changes, which act on a yearly basis, or changes in price on the world market. Changes in global oil price are already reflected in the monthly average sales price reported by producers. I also model production with and without the inclusion of an additional regressor to indicate when a lease permanently stops producing. The decision by an operator to stop producing from the lease is governed by the price of oil after royalty payments and taxes, as well as the producing age of the lease and the cost factors captured by the lease fixed effect. It is therefore conditionally exogenous in this framework. Including the indicator controls for the production lost due to the operators' decisions to abandon fields. The resulting production changes will then only be the result of changes in the rate of production from leases that keep producing.[20] Lease shut-in decisions are modeled separately. My production model is as follows:

$$lnQ_{it} = \beta_0 + \beta_i lnP_{it} + \gamma_{1i}Age_{it} + \gamma_{2i}Age_{it}^2 + \eta_i ShutIn_{it} + \delta_t + \alpha_i + \epsilon_{it} \tag{7}$$

Q_{it} is the average daily production from lease i in month t; P_{it} is the actual sale price of oil from each lease each month net of royalty payments and severance taxes; Age_{it} reflects the number of months since first production from each lease to capture, along with its square, the effect of depletion; $ShutIn_{it}$ is an indicator that changes from 0 to 1 when a lease permanently stops producing for the remainder of the sample; δ_t is a time trend to capture technological advances; α_i represents lease fixed effects and picks up production cost heterogeneity reflected in characteristics such as the depth of the reservoir, gravity of the oil (which is also partially reflected in the price), recovery method, and transport to market.

In addition to the specification in equation (7), monthly price is replaced with lagged price. Prices with up to twelve monthly lags are used, as well as three-month and twelve-month average price. If expectations of the future price, which could be driving production decisions, are based on prices in the previous months, lagged price may be more appropriate. Furthermore, producers may have entered into crude oil supply contracts in previous months based on prices in those months, and it could be those contractual obligations driving supply. Total contract supply can also be met from production from other properties leased by the producer, whether on private, state, federal, or Indian lands, however, it is not possible to look at aggregate production decisions for each operator with the present data set.

[20] These production decisions could include reducing rates deliberately from all wells on the lease, or taking some wells out of service while allowing others to maintain production.

Issues estimating appropriate standard errors arise from the likely presence of a high degree of serially correlation, borne out by the results as indicated in Figures 8 and 9 and discussed in section 2.6. Driskoll-Kraay standard errors, robust to both spatial and serial error autocorrelation, are therefore estimated. The Levin et al. (2002) test of unit roots in panel data fails to reject the null of nonstationarity of oil production, whereas a unit root test on the first-difference strongly rejects the presence of unit roots in the panels. I therefore also estimate a first-difference version of equation (7).

2.4.2 Error Correction Model

Given the potential for nonstationarity and cointegration, a dynamic specification is also used to estimate the long-run production response to changes in price. Westerlund's 2007 error correction-based tests of cointegration strongly reject the no cointegration null. To account for these issues I employ a dynamic fixed effects estimator (DFE) for nonstationary panels with both large cross-section and time dimensions, as described in Cameron and Trivedi (2009).

The error correction reparameterization of the autoregressive distributed lag dynamic panel specification of model (7) is as follows:[21]

$$\Delta q_{it} = \phi_i(q_{i,t-1} - \theta_{0i} - \theta_{1i} p_{it}) + \delta_{1i}\Delta p_{it} + \gamma_{1i} Age_{it} + \gamma_{2i} Age_{it}^2 + \eta_i ShutIn_{it} + \delta_t + \epsilon_{it} \quad (8)$$

where the parameter of interest is θ_{1i}, the long-run price elasticity of supply. For a complete description of the parameters, see Pesaran and Smith (1995) or Blackburne and Frank (2007).

2.4.3 Lease Survival Analysis

To assess the effectiveness of the Stripper Oil Well Property Royalty Rate Reduction Program, I conduct a survival analysis to model the time to a lease shut-in event. If the program was effective, then participating leases should have a lower probability of being

[21] For simplicity, an error correction reparameterization of an ARDL(1,1) panel specification is shown, but it could be generalized to an ARDL(p,q). Lower case variables in the model refer to their natural logarithms.

shut-in at a particular age, controlling for price and other relevant factors. A Cox proportional hazard model is estimated using the combined data set of both stripper well and non-stripper well leases. The probability that the lease will survive at a particular producing age, given that it is still producing, is theorized to be a function of the price of oil, the volume of production, an interaction between the two, and an indicator of whether or not the lease participated in the royalty reduction program. Given that the effect of price on lease survival is highly non-linear, as discussed in Section 2.6.3, a fourth-order polynomial in price is used. Further, hazard proportionalities of the higher-order functions of price are rejected, so the interactions of these terms with the log of time (or in this case, the age of the lease) are also included.

Production costs are also critical in determining the producing life of a lease. However, this data is not presently available to researchers. Given the approximate normality of the estimated fixed effects used to account for production cost heterogeneity in the models in the preceding sections (see Figures 8 and 9), results from the hazard model may still be informative about the success of the program on average.

2.5 Data

The data set compiled for this study was provided on a confidential basis from two federal agencies within the Department of the Interior. Federal lease history data was provided by the BLM. Although the components of this data can be accessed publicly through their LR2000 web portal,[22] it is not possible to compile large data sets in a convenient format in this manner. The lease history data includes the location of each federal lease, the date of first production, the base royalty rate, and the date of all changes in the royalty rate as a result of participation in the stripper well royalty reduction program.

Data on monthly lease production was provided by the Office of Natural Resources Revenue (ONRR) for all federal leases that reported production between January 1996 and May 2011. Data for the subset of producing leases that qualified for the stripper well royalty reduction program was supplied from January 1990 onwards. The data includes the

[22] `http://www.blm.gov/landandresourcesreports/rptapp/menu.cfm?appCd=2`

actual volume of oil sold each month by each lessee, the total value of the oil sold, and the total amount of royalty paid to ONRR. From this information, daily average production, the actual royalty rate, and the price net of royalties can be calculated. The price net of severance taxes is then calculated from state severance rate data gathered from state resource or tax agencies.[23]

For months when leases did not report oil sales, prices were not supplied and had to be estimated for the particular type of oil that the lease typically produced. This was accomplished by calculating the average ratio between the actual reported sales price of the lease's oil for the months when sales did occur and the monthly price of WTI. This average ratio was then multiplied by the price of WTI to estimate what the lease's oil could have commanded on the market. Months with zero production were changed to a small positive number to allow taking the natural log of the data.

For smaller producers, especially producers that rely on tanker trucks to move oil from the field to refineries, inventory may be allowed to build until enough oil is collected to fill a truck. Inspecting the sales history data, it is evident that for many leases several months can pass between sales of several hundred barrels. In these situations, simply recording a small positive number for months when sales did not occur would not reflect reality. I try four methods to account for this issue. First, the cumulative production for months between sales is averaged over that period. If no sales were made for three months, and then in the fourth month a sales volume was recorded, Q_{it} for each of those four months would be the total sales in month four divided by four. I refer to this as the production averaging method. Second, I aggregate the sales to quarterly sales data and use the average quarterly price. Third, I aggregate to yearly sales and average yearly price. Fourth, I use data for only those leases which report sales for more than 150, 200, and 250 months of the sample period for stripper well leases, for which there are 257 total months (the corresponding numbers for non-stripper well leases are 100, 150, and 180 out of 185 months). Shut in months are counted as months for which there is a sales report since the producer has decided to produce no oil. Production averages, as in the first method, are then used for

[23]Severance taxes are not assessed on the royalty interests due federal or state governments or to Indian tribes.

the small percentage of months without a reported sale.

Summary Statistics for federal leases which qualified for the Stripper Oil Well Property Royalty Rate Reduction Program are supplied in Table 9 in Appendix D for leases with more than 200 reported sales (out of 257 total months). Summary Statistics for non-stripper well leases with more than 150 reported sales (out of 185 total months) can be found in Table 24 in Appendix E. On average, stripper well leases have been producing longer than those that did not qualify for the program; in May of 2011, their respective ages were 480 and 374 months. This is no surprise as stripper well leases are generally nearing the end of their producing lives and older leases would be most likely to qualify for the program. Average monthly production is slightly higher on stripper well leases than on non-stripper well leases; 36.6 and 34.1 BPD respectively. Qualifying as a stripper well lease is based on average per well production. Some very large leases with over a thousand wells could have a very high total daily production and low per well average, thereby qualifying them for royalty relief. This may lead to the slightly higher lease average production observed if large leases were more likely to be included in the stripper well sample. An example would be if older BLM auction parcels were larger in size than more recent auctions. In addition, the near record prices after the end of the stripper well royalty reduction program may have caused leases to start producing in areas with only small volumes of expected oil that could previously not be economically recovered. These leases would produce at low rates which would result in a reduction in average production for the pool of leases that weren't involved in the stripper well lease program. Approximately 25% of the non-stripper well leases started production after the royalty reduction program had ended. Excluding these leases increases the average production of non-stripper well leases to 38.3 BPD. The average royalty paid on the oil from stripper well leases was 9.0% compared to roughly 12% royalty for all other federal leases. Heavier oil wells were also afforded some royalty relief and are included in the sample of non-stripper well leases, making the average slightly lower than the standard $12\frac{1}{2}\%$ royalty for most federal leases. The average price net of royalties and tax for stripper well leases was \$29.70 (May 2011 dollars) compared to \$38.70 for non-stripper well leases. The stripper well data includes the six additional years from

1990 to 1995, when prices were lower on average, making the overall average price in the stripper well sample lower.

Figures 6 and 7 are plots of production and price for a sample of stripper well and non-stripper well leases, respectively. As is apparent from the plots of oil price, for this sample of leases the prices across leases tend to follow each other closely. The period during which royalty reductions were in effect is noted in Figure 6. The program was implemented during a period of low prices and discontinued after prices started increasing sharply in 2005.

2.6 Results

2.6.1 Static Panel Regressions

Tables 10 through 20 in Appendix C are the estimates from equation (7) using data from stripper well leases. Tables 25 through 35 in Appendix D are estimates from equation (7) using data from non-stripper well leases. Tables 10 and 25 are the pooled OLS, panel between effects, and panel random effects estimates for each group of leases counting no sales report for a month as zero lease production. Tables 11 and 26 report corresponding estimates from fixed effects regressions. Hausman tests strongly reject random effects in favor of fixed effects for both pools of leases. While elasticity estimates are large and significant with robust standard errors clustered at the lease-level, assumptions that no sales report for a month indicates zero production are clearly not warranted. Further evidence challenging this assumption is given by the very low predictive power these estimates provide. Also of note are the lack of significance when Driscoll-Kraay standard errors are estimated. Wooldridge's (2002) test of serial correlation in panel data strongly rejects the null of zero first order autocorrelation for both stripper and non-stripper well qualified properties. By accounting for the high degree of serial correlation in the residuals, highly significant elasticity estimates become insignificant.

Tables 12 and 27 provide GLS fixed effects estimates when production averaging is used to account for periodic sales of cumulative production. For the pool of leases qualified as stripper well properties, significant elasticities of approximately 0.24 are obtained. Elasticity estimates are small and insignificant for all other leases. In both cases, the predictive power of the models are vastly improved. Using lagged prices provides similar results in both cases.

Tables 13 and 28 provide GLS fixed effects estimates when production is aggregated quarterly and the average quarterly price is used. Tables 14 and 29 provide GLS fixed effects estimates when production is aggregated annually and average annual price is used. While non-stripper well lease elasticities are still small and insignificant, those for the stripper well leases are again significant and inflated. Explanatory power is quite low in all four cases compared to the model using production averaging. The remainder of the results therefore use production averaging to account for producing months without a sales report.

Using production averaging results in a loss of variation in monthly production rates and also requires the use of estimates of monthly prices for the months when sales did not occur. This could impact estimates of the responsiveness of producers to changes in price. To overcome this, I next look at those leases for which sales are reported with varying frequencies. I look at leases with sales in more than roughly 60% of the sample period, less than 60% of the sample period, more than 80%, and more than 95% of the sample period. For stripper well leases the results are presented in Tables 15, 16, 17, and 18, respectively. The corresponding results for non-stripper well leases are presented in Tables 30, 31, 32, and 33, respectively.

For stripper well leases, the higher the percentage of months with a reported sale, the smaller and less significant is the elasticity estimate. Estimates range from 0.14 with a p-value less than 0.001 (more than 60% of months with sales reports), down to 0.03 significant at 10% (more than 95% of months with sales reports). Results also vary slightly by using lagged prices, and the estimates generally lose considerable explanatory power by using average twelve month prices. When only leases with sales reports in less than roughly 60% of the sample period are used, which generally indicates that the lease produces small amounts of oil daily, the estimated elasticity is quite high compared to other stripper well leases (between 0.44 and 0.56 depending on which price is used). As the number of months between sales varies between two months to over a year, it is unclear to which price producers are responding as they may not know in advance the ultimate sales price, absent a supply contract with a refiner. Given the length of time between sales, it is more likely that elasticities for these leases are reflecting either the shutting down of wells on the lease over time, potentially drilling new wells, or implementing enhanced recovery methods to increase production. These are changes that could easily occur between sales reports.

Results with varying frequencies of sales reports are not as clear for non-stripper well leases. For leases with sales reported in less than roughly 60% of the months, elasticities are small and positive, yet insignificant given relatively large standard errors. For leases with sales reported in more than 60% of the months, elasticities are small, negative, and insignificant. For leases with sales in more than both 80% and 95% of the months, results are small, negative, and marginally significant. When lagged prices are used, however, elasticities are closer to zero and insignificant. Again, using the average of the previous twelve-month price reduces the explanatory power of the model considerably.

To check the robustness of results with prices over longer time horizons, I also run regressions with three-month and twelve-month rolling average production and price for data with varying frequencies of sales reports. This smooths out variations based on short-term production decisions that may be related to, for example, issues with oil storage capacity or temporary refining bottlenecks in certain regions. These results are presented in Tables 19 and 20 for stripper well leases and in Tables 34 and 35 for non-stripper well leases. Results are generally consistent with the results described previously. However, the results of dynamic panel regressions are a more effective analysis for production responses over longer time horizons with panel data consisting of large time and cross section components. These results are discussed in Section 2.6.2.

Similarly, the results of first-difference estimation, found in Tables 21 and 36, largely conform to the elasticity estimates described above. There are two main exceptions. The elasticity of supply for stripper well leases with less than 150 sales is 0.086, significantly less than the fixed effects estimates. Also, the elasticity estimates for non-stripper well leases are more significant and tend to be smaller negative numbers.

2.6.2 Dynamic Panel Regressions

Estimates from the dynamic fixed effects model yield long-run results which are largely consistent with the static analysis above. Estimates of equation (8) for stripper well leases are presented in Table 22 and the corresponding non-stripper well lease results can be found in Table 37. For stripper well leases with regularly reported production, long-run elasticities range from zero to 0.06. Those stripper well leases without regularly reported sales have a long-run elasticity of approximately 0.62, slightly higher than the static panel estimates. Non-stripper well leases with regularly reported sales have small negative, insignificant long-run elasticities. Unlike the static panel analysis, however, non-stripper well leases without regularly reported sales have a marginally significant long-run elasticity of approximately 0.06. Differenced log prices up to nine lags are used; an ARDL(1,9) model. The inclusion of additional lags are not informative.

2.6.3 Lease Survival Analysis

The results of a log-rank test of equality of the survivor functions for the stripper well and non-stripper well program leases strongly rejects equality. A plot of the Kaplan-Meier survival estimates shown in Figure 10 confirms this and provides evidence that participating in the program extended the producing lives of leases. Table 23 shows the results from the Cox proportional hazard model. A test of the proportional hazards assumption based on the scaled Schoenfeld residuals rejects proportional hazard of the higher-order price terms, indicating that interactions with lease age should be included (the log of lease age is used in the interaction variable). From this final model, the estimated hazard ratio of the stripper well indicator shows that participation in the Stripper Oil Well Property Royalty Rate Reduction Program lead to a decrease in the probability that a lease was shut-in of approximately 15%, controlling for price, the rate of production in the month prior to shut-in, an interaction between price and production, and interactions of the higher-order price variables and time.

This result should be viewed with caution, however, given the lack of data for marginal properties that may have been shut-in prior to the start of the program. Data for leases that participated in the royalty reduction program was supplied from 1990 onward, about three years prior to its start. Data for all leases was supplied beginning in 1996, about 3 years after the program started. Therefore, a marginal lease that shut-in between January 1990 and the start of the program at the end of 1992 when oil prices were relatively low, will not be included in the data. Since prices remained relatively high after the program was discontinued, and presumably all eligible properties participated, the current data does not provide the type of counterfactual that would be necessary to verify these results. This can be rectified with assistance from the ONRR to generate the required data.

The coefficients on price up to a fourth-order polynomial, apart from the quadratic term, were significant in this model. At higher price levels, particularly above $60 per barrel, changes in price lead to an exponentially smaller probability that a lease will stop producing than at lower levels, as inferred from the coefficients of the price terms. Further, the significance of the time varying coefficients indicates that the reduction in probability a lease will be shut-in when a price increases will be smaller as a lease has been producing longer. This would make sense if costs are expected to increase over time as the producing life of the lease increases. Larger price increases would therefore be necessary to prevent shut-in.

2.7 Conclusions and Extensions

Based on this analysis, I can conclude that production from the class of leases that qualified as federal stripper oil well leases is relatively more elastic than production from non-stripper well leases, regardless of the specification. However, all but the most marginal of producing leases—those leases that participated in the program and did not report regular sales—have estimated price elasticities of supply not significantly different than zero in most cases. Zero supply elasticities should not be taken as a lack of a competitive market structure. This data set captures only the price response once operators start producing. Decisions to invest in exploration and to bring new producing leases online in response to high prices is not captured by the analysis conducted here. It is at this extensive margin that the overall supply elasticity will be determined. By looking at leasing and exploratory drilling decisions as a function of state versus federal fiscal regimes (royalty, taxes, lease bonus bids, and lease rental rates), more insight can be gained about producer responses.

These results have important implications for policy related to oil taxation. First, by reducing production royalties for the most marginal of leases during periods of lower prices, producers can be expected to respond by making the additional investments required to extend the producing life of reservoirs. While this result is not unexpected, this analysis provides evidence using a unique data set and lends justification to the Stripper Oil Well Property Royalty Rate Reduction Program. Second, and perhaps most importantly from a fiscal standpoint, higher production royalties do not appear to deter production for the bulk of federal leases. This suggests that additional revenue could be raised by increasing royalties on federal lands from $12\frac{1}{2}\%$ for new leases, for example to a level more in line with those required by state governments from production on state lands (generally $16\frac{2}{3}\%$). Combining these two results suggests that a higher initial royalty rate, in concert with a lower rate for marginal leases during periods of lower relative oil price, can be an effective means to raise revenue without negatively impacting producing life. Further analysis, as described above, must be conducted to see what impact higher initial production royalties will have at the extensive margin.

In an attempt to address estimation issues related to the unitization of leases, using correlations between lease production within each state can be explored as a means to determine which federal leases belong to units and should therefore have their productions combined. Highly correlated production rates approaching one for leases within the same county, or in neighboring counties, signify that they stem from a common production divided up amongst the leases that form the unit. Preliminary results suggest production from a number of federal leases within the same counties do tend to be highly correlated, suggesting that they are part of units. If a sample of these highly correlated leases are tracked down to see if they are actually part of a common unit, this correlation approach could be confirmed. The results can be rechecked after the highly correlated lease productions are added together. However, given the lack of significance for most results, this is not likely to change the core findings of this study.

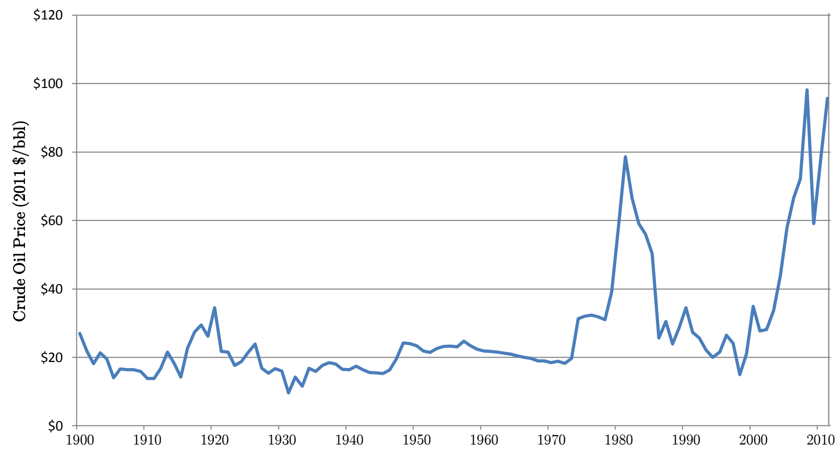
Appendix C: Schedule of Royalty Rate Reductions

Average Number of Barrels Reduced of Oil Produced Per Well Per Day	Royalty Rate Percentage
0	0.5
1	1.3
2	2.1
3	2.9
4	3.7
5	4.5
6	5.3
7	6.1
8	6.9
9	7.7
10	8.5
11	9.3
12	10.1
13	10.9
14	11.7

Source: LaRouche 2001

Appendix D: Figures

Figure 4 – *U.S. Wellhead First Purchase Crude Oil Price 1900 - 2011*

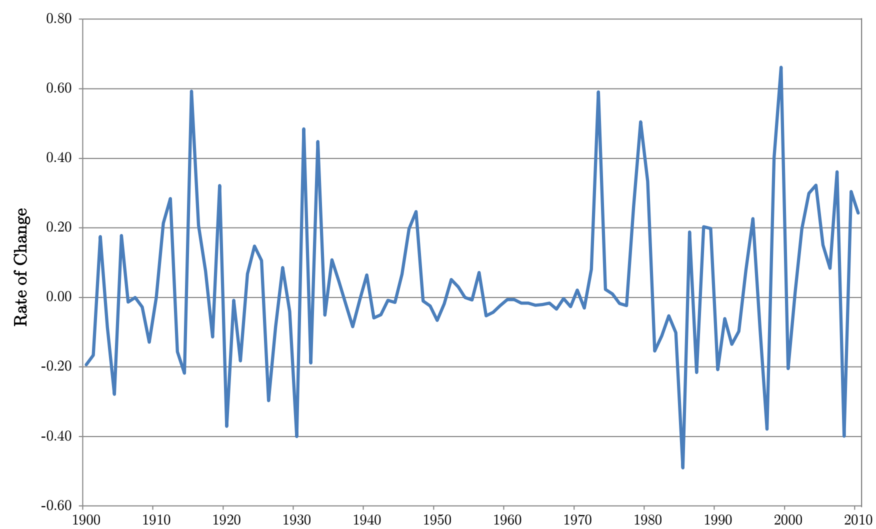


U.S. Wellhead First Purchase Price

Data sources: U.S. Energy Information Administration at http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx

Bureau of Labor Statistics at http://www.bls.gov/cpi/tables.htm

Figure 5 – *Rate of Change of U.S. Oil Price, 1900 - 2011*



U.S. Wellhead First Purchase Price

Data sources: U.S. Energy Information Administration at http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx

Bureau of Labor Statistics at http://www.bls.gov/cpi/tables.htm

Figure 6 – *Log of Avg Daily Production and Price net of Royalty and Severance Tax (May 2011 \$) for a sample of Stripper Well Leases (Jan. 1990 – May 2011)*

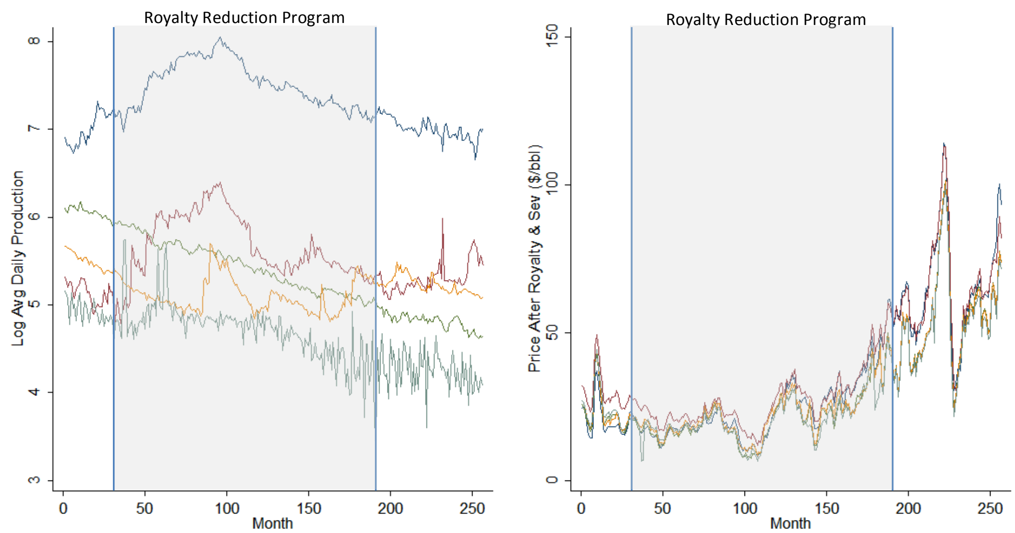


Figure 7 – *Log of Avg Daily Production and Price net of Royalty and Severance Tax (May 2011 \$) for a sample of Non–Stripper Well Leases (Jan. 1996 – May 2011)*

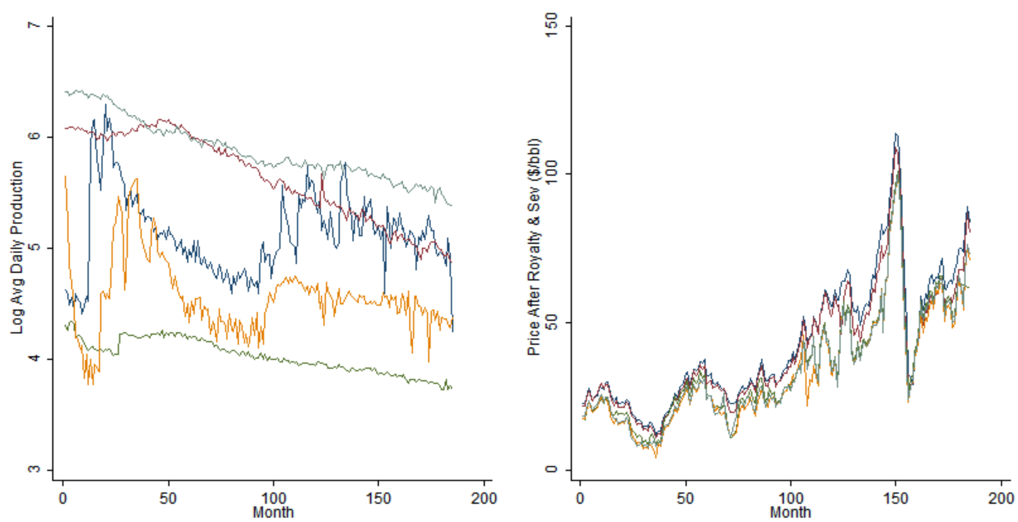
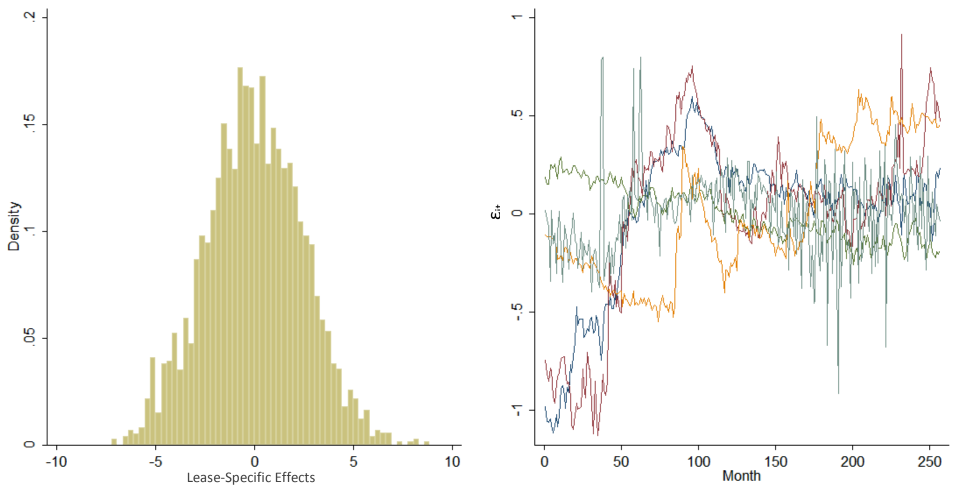
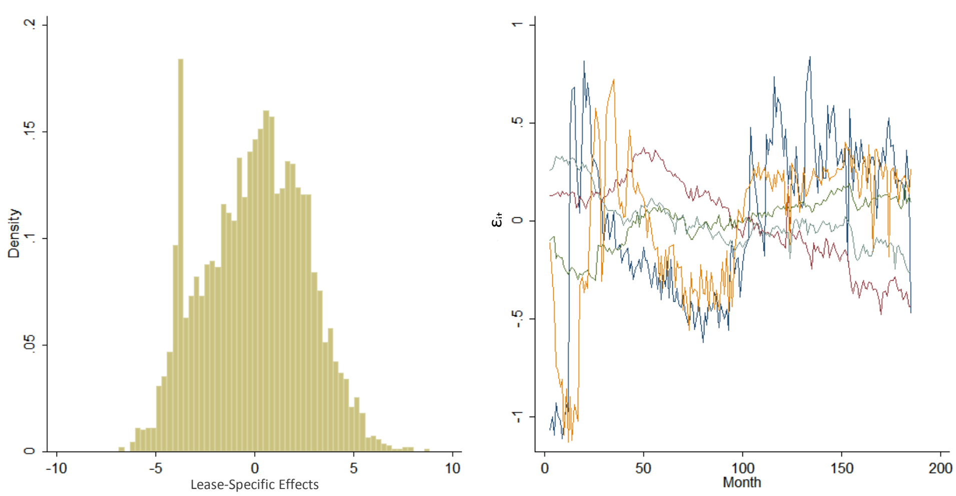


Figure 8 – *Estimated density of the estimated lease-specific effects,* $\{\hat{\alpha}_i\}_{i=1}^{2,657}$ *(left), and time series plots of* $\{\hat{\epsilon_{it}}\}$ *from a sample of five leases (right)*



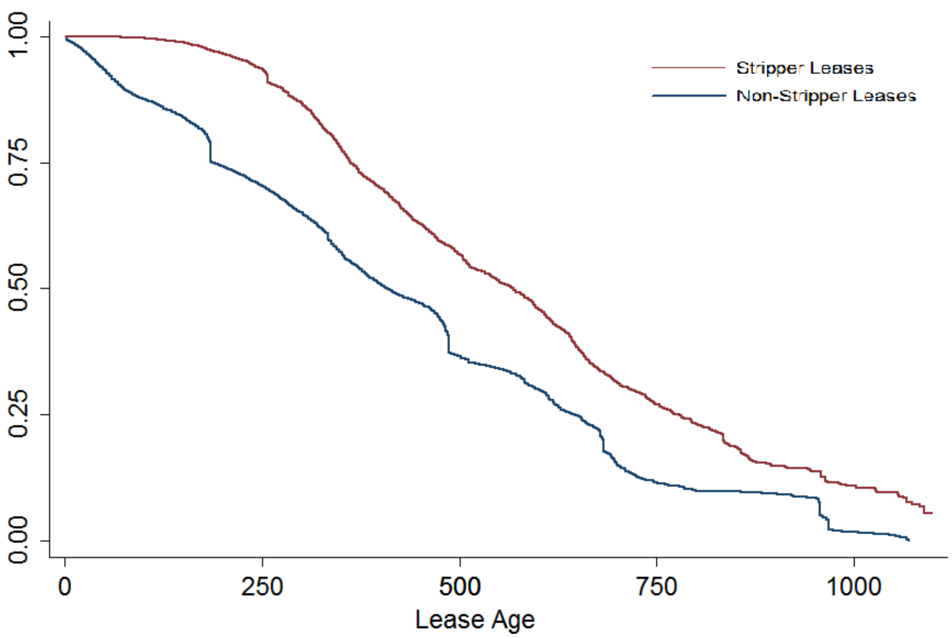
These results apply to fixed effects regressions of equation (7) for the 2,657 stripper well leases with > 200 reported sales.

Figure 9 – *Estimated density of the estimated lease-specific effects,* $\{\hat{\alpha}_i\}_{i=1}^{4,116}$ *(left), and time series plots of* $\{\hat{\epsilon_{it}}\}$ *from a sample of five leases (right)*



These results apply to fixed effects regressions of equation (7) for the 4,116 non–stripper well leases with > 150 reported sales.

Figure 10 – *Kaplan–Meier Survival Estimates for Stripper Well and Non–Stripper Well Leases*



Appendix E: Stripper Oil Well Lease Summary Statistics and Regression Results

Table 9 – *Summary Statistics for Stripper Well Leases with more than 200 Reported Sales (Jan. 1990 – May 2011)*

Variable	**Mean**	**Std. Dev.**	**N**
Avg Daily Prod, BPD	36.60	138	687,512
Age in May 2011	480	197	2,657
Shut-In Status	0.038	0.191	687,512
Price Net of Royalty	29.70	20.40	681,344
Royalty Rate	0.090	0.041	687,512
Total Quarterly Production	3,347	12,474	227,404
Avg Quarterly Oil Price	29.29	19.70	225,364
Total Yearly Production	13,400	49,412	56,205
Avg Yearly Oil Price	28.80	18.25	55,701
Number of Leases		2,657	

Table 10 – *Oil Production Regression Results for the Stripper Well Lease Full Sample: OLS, BE, and RE Model Comparisons Treating no Sales Report for a Month as Zero Lease Production*

	OLS	**BE**	**RE**
Log Price, ϵ_s	0.267***	−3.470***	0.510***
	(0.054)	(0.556)	(0.031)
Months Producing	−2.5e-04	−1.9e-06	1.4e-03*
	(8.4e-04)	(1.2e-03)	(6.3e-04)
Months Producing Sqrd	4.9e-06***	4.5e-06**	2.9e-06***
	(9.5e-07)	(1.4e-06)	(6.7e-07)
Shut-In Status	−8.624***	−14.569***	−4.406***
	(0.075)	(0.417)	(0.129)
Trend	−0.007***	0.039***	−0.011***
	(0.000)	(0.008)	(0.000)
Constant	−2.06***	5.05**	−2.81***
	(0.23)	(1.96)	(0.16)
No. Obs.	1,201,152	1,201,152	1,201,152
No. Groups		4,731	4,731
F	3,937	303	
R^2	0.12	0.24	
R^2_o		0.056	0.11
R^2_b		0.24	0.16
R^2_w		0.0022	0.052
σ_α			3.67
σ_e			3.85
ρ (Var. fraction due to α_i)			0.48

Robust standard errors clustered at the lease level in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11 – *Oil Production Regression Results for the Stripper Well Lease Full Sample: Panel Fixed Effects Model Treating no Sales Report for a Month as Zero Lease Production*

	GLS, FE	**FE, DK**	**FE–REGAR**
Log Price, ϵ_s	0.511***	0.511***	0.527***
	(0.031)	(0.122)	(0.013)
Months Producing	−9.9e-03***	−9.9e-03***	−9.7e-03***
	(5.9e-04)	(1.3e-03)	(1.4e-04)
Months Producing Sqrd	2.9e-06***	2.9e-06***	2.7e-06***
	(6.8e-07)	(4.3e-07)	(1.6e-07)
Shut-In Status	−4.4***	−4.4***	−4.7***
	(0.13)	(0.33)	(0.03)
Constant	−0.26+	−0.26	−0.36***
	(0.13)	(0.23)	(0.03)
No. Obs.	1,201,152	1,201,152	1,196,421
No. Groups	4,731	4,731	4,731
F	538	467	11,550
R^2_w	0.052	0.052	0.037
σ_α	4.48	4.48	
σ_e	3.85	3.85	
ρ (Var. fraction due to α_i)	0.57	0.57	

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with “Months Producing” variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 100 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12 – *Oil Production Regression Results for the Stripper Well Lease Full Sample: Panel Fixed Effects Model Using Production Averaging to Account for Intermittent Sales of Cumulative Production*

	GLS, FE	**FE, DK**	**FE, DK**
Log Price, ϵ_s	0.240***	0.240***	
	(0.014)	(0.045)	
Log 2-Month Lag Price, ϵ_s			0.232***
			(0.047)
Months Producing	−4.9e-03***	−4.9e-03***	−4.8e-03***
	(2.8e-04)	(4.4e-04)	(4.5e-04)
Months Producing Sqrd	7.1e-07*	7.1e-07***	6.9e-07***
	(3.1e-07)	(1.2e-07)	(1.2e-07)
Shut-In Status	−13.0***	−13.0***	−13.0***
	(0.07)	(0.40)	(0.40)
Constant	5.17***	5.17***	5.16***
	(0.06)	(0.09)	(0.08)
No. Obs.	1,201,243	1,201,243	1,192,347
No. Groups	4,731	4,731	4,731
F	10,969	904	1,000
R^2_w	0.79	0.79	0.79
σ_α	2.31	2.31	2.30
σ_e	1.14	1.14	1.14
ρ (Var. fraction due to α_i)	0.80	0.80	0.80

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 100 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13 – *Oil Production Regression Results for the Stripper Well Lease Full Sample: Panel Fixed Effects Model Using Quarterly Aggregation*

	GLS, FE	**FE, DK**	**FE, DK**
Log Avg Quarterly Price, ϵ_s	0.488***	0.488***	
	(0.032)	(0.089)	
Log 2-Quarter Lag Price, ϵ_s			0.355**
			(0.129)
Months Producing	−9.2e-03***	−9.2e-03***	−8.6e-03***
	(5.4e-04)	(9.4e-04)	(1.2e-03)
Months Producing Sqrd	2.9e-06***	2.9e-06***	3.0e-06***
	(6.1e-07)	(2.7e-07)	(2.9e-07)
Shut-In Status	−5.4***	−5.4***	−5.3***
	(0.14)	(0.34)	(0.33)
Constant	5.38***	5.38***	5.62***
	(0.12)	(0.23)	(0.23)
No. Obs.	397,348	397,348	388,408
No. Groups	4,731	4,731	4,731
F	603	72	72
R^2_w	0.091	0.091	0.089
σ_α	3.72	3.72	3.67
σ_e	3.08	3.08	3.08
ρ (Var. fraction due to α_i)	0.59	0.59	0.59

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 25 lags)

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14 – *Oil Production Regression Results for the Stripper Well Lease Full Sample: Panel Fixed Effects Model Using Yearly Aggregation*

	GLS, FE	**FE, DK**	**FE, DK**
Log Avg Yearly Price, ϵ_s	0.316***	0.316***	
	(0.036)	(0.058)	
Log Avg 2-Year Price, ϵ_s			0.514***
			(0.073)
Months Producing	−6.6e-03***	−6.6e-03***	−7.1e-03***
	(4.9e-04)	(6.6e-04)	(4.6e-04)
Months Producing Sqrd	2.0e-06***	2.0e-06***	1.5e-06***
	(5.4e-07)	(1.2e-07)	(1.6e-07)
Shut-In Status	−5.4***	−5.4***	−5.2***
	(0.14)	(0.66)	(0.63)
Constant	7.42***	7.42***	7.03***
	(0.12)	(0.10)	(0.17)
No. Obs.	98,267	98,267	89,212
No. Groups	4,731	4,731	4,731
F	557	135	229
R_w^2	0.15	0.15	0.14
σ_α	2.88	2.88	3.03
σ_e	2.30	2.30	2.27
ρ (Var. fraction due to α_i)	0.61	0.61	0.64

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 10 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15 – *Oil Production Regression Results for Stripper Well Leases With More Than 150 Reported Sales: Panel Fixed Effects Model Averaging Production for Months with no Sales Report*

	GLS, FE	**FE, DK**	**FE, DK**	**FE, DK**	**FE, DK**
Log Price, ϵ_s	0.140***	0.140***			
	(0.015)	(0.025)			
Log 2-Month Lag Price, ϵ_s			0.140***		
			(0.027)		
Log 4-Month Lag Price, ϵ_s				0.128***	
				(0.032)	
Log Avg 12 Month Lag Price, ϵ_s					0.248***
					(0.074)
Months Producing	−4.1e-03***	−4.1e-03***	−4.1e-03***	−4.0e-03***	−7.4e-03***
	(3.0e-04)	(2.1e-04)	(2.1e-04)	(2.2e-04)	(7.8e-04)
Months Producing Sqrd	2.0e-07	2.0e-07	1.9e-07	2.0e-07	1.5e-06***
	(3.3e-07)	(1.8e-07)	(1.9e-07)	(2.0e-07)	(2.1e-07)
Shut-In Status	−14***	−14***	−14***	−14***	−6.6***
	(0.085)	(0.34)	(0.34)	(0.34)	(0.88)
Constant	5.96***	5.96***	5.95***	5.96***	2.08***
	(0.07)	(0.05)	(0.05)	(0.05)	(0.13)
No. Obs.	871,394	871,394	864,860	858,311	831,981
No. Groups	3,416	3,416	3,416	3,416	3,416
F	8,127	1,773	1,958	2,130	525
R^2_w	0.78	0.78	0.78	0.78	0.090
σ_α	2.18	2.18	2.18	2.18	3.04
σ_e	1.05	1.05	1.05	1.05	3.29
ρ (Var. fraction due to α_i)	0.81	0.81	0.81	0.81	0.46

Robust standard errors clustered at the lease level in parentheses
Trend omitted due to collinearity with "Months Producing" variable
FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 100 lags)
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 16 – *Oil Production Regression Results for Stripper Well Leases With Less Than 150 Reported Sales: Panel Fixed Effects Model Averaging Production for Months with no Sales Report*

	GLS, FE	**DK, FE**	**DK, FE**	**DK, FE**	**DK, FE**
Log Price, ϵ_s	0.503***	0.503***			
	(0.032)	(0.094)			
Log 2-Month Lag Price, ϵ_s			0.469***		
			(0.093)		
Log 4-Month Lag Price, ϵ_s				0.439***	
				(0.100)	
Log Avg 12 Month Lag Price, ϵ_s					0.563***
					(0.096)
Months Producing	−7.4e-03***	−7.4e-03***	−7.1e-03***	−6.9e-03***	−7.1e-03***
	(6.0e-04)	(9.4e-04)	(9.6e-04)	(1.0e-03)	(8.3e-04)
Months Producing Sqrd	3.2e-06***	3.2e-06***	3.2e-06***	3.2e-06***	3.0e-06***
	(8.0e-07)	(2.2e-07)	(2.1e-07)	(2.0e-07)	(1.6e-07)
Shut-In Status	−12***	−12***	−12***	−12***	−12***
	(0.083)	(0.16)	(0.16)	(0.16)	(0.18)
Constant	3.06***	3.06***	3.09***	3.12***	2.79***
	(0.13)	(0.17)	(0.16)	(0.16)	(0.15)
No. Obs.	327,158	327,158	324,814	322,457	312,892
No. Groups	1,304	1,304	1,304	1,304	1,304
F	7,887	5,102	5,153	4,888	6,686
R^2_w	0.80	0.80	0.80	0.80	0.81
σ_α	1.68	1.68	1.66	1.65	1.68
σ_e	1.30	1.30	1.30	1.30	1.29
ρ (Var. fraction due to α_i)	0.63	0.63	0.62	0.62	0.63

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 100 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 17 – *Oil Production Regression Results for Stripper Well Leases With More Than 200 Reported Sales: Panel Fixed Effects Model Averaging Production for Months with no Sales Report*

	GLS, FE	**FE, DK**	**FE, DK**	**FE, DK**	**FE, DK**
Log Price, ϵ_s	0.069^{***}	0.069^{***}			
	(0.015)	(0.021)			
Log 2-Month Lag Price, ϵ_s			0.071^{**}		
			(0.025)		
Log 4-Month Lag Price, ϵ_s				0.059^{+}	
				(0.031)	
Log Avg 12 Month Lag Price, ϵ_s					0.029
					(0.094)
Months Producing	$-3.6\text{e-}03^{***}$	$-3.6\text{e-}03^{***}$	$-3.6\text{e-}03^{***}$	$-3.5\text{e-}03^{***}$	$-4.2\text{e-}03^{***}$
	(3.2e-04)	(2.4e-04)	(2.4e-04)	(2.3e-04)	(6.9e-04)
Months Producing Sqrd	−7.7e-09	−7.7e-09	−1.1e-08	1.4e-08	−3.2e-07
	(3.4e-07)	(2.4e-07)	(2.5e-07)	(2.6e-07)	(3.3e-07)
Shut-In Status	-14^{***}	-14^{***}	-14^{***}	-14^{***}	-8.3^{***}
	(0.11)	(0.43)	(0.43)	(0.43)	(0.80)
Constant	6.36^{***}	6.36^{***}	6.35^{***}	6.37^{***}	2.88^{***}
	(0.08)	(0.05)	(0.06)	(0.07)	(0.11)
No. Obs.	681,344	681,344	676,160	670,965	650,101
No. Groups	2,657	2,657	2,657	2,657	2,657
F	4,503	3,281	3,104	3,368	447
R^2_w	0.78	0.78	0.78	0.78	0.15
σ_α	2.12	2.12	2.13	2.12	2.46
σ_e	0.93	0.93	0.92	0.92	2.53
ρ (Var. fraction due to α_i)	0.84	0.84	0.84	0.84	0.49

Robust standard errors clustered at the lease level in parentheses
Trend omitted due to collinearity with "Months Producing" variable
FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 100 lags)
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 18 – *Oil Production Regression Results for Stripper Well Leases With More Than 250 Reported Sales: Panel Fixed Effects Model Averaging Production for Months with no Sales Report*

	GLS, FE	**FE, DK**	**FE, DK**	**FE, DK**	**FE, DK**
Log Price, ϵ_s	0.031^{+}	0.031^{+}			
	(0.017)	(0.018)			
Log 2-Month Lag Price, ϵ_s			0.034^{*}		
			(0.016)		
Log 4-Month Lag Price, ϵ_s				0.024	
				(0.015)	
Log Avg 12 Month Lag Price, ϵ_s					0.035^{***}
					(0.008)
Months Producing	$-2.9\text{e-}03^{***}$	$-2.9\text{e-}03^{***}$	$-2.9\text{e-}03^{***}$	$-2.9\text{e-}03^{***}$	$-3.0\text{e-}03^{***}$
	(4.0e-04)	(3.4e-04)	(3.4e-04)	(3.4e-04)	(4.5e-04)
Months Producing Sqrd	−2.1e-07	−2.1e-07	−1.9e-07	−1.6e-07	−1.2e-07
	(4.0e-07)	(3.6e-07)	(3.6e-07)	(3.6e-07)	(4.5e-07)
Shut-In Status	-14^{***}	-14^{***}	-14^{***}	-14^{***}	-11^{***}
	(0.26)	(0.63)	(0.62)	(0.62)	(0.16)
Constant	6.81^{***}	6.81^{***}	6.80^{***}	6.82^{***}	3.36^{***}
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
No. Obs.	372,647	372,647	369,750	366,850	355,240
No. Groups	1,450	1,450	1,450	1,450	1,450
F	907	1,690	1,645	1,875	4,947
R^2_w	0.68	0.68	0.68	0.68	0.41
σ_α	2.19	2.19	2.20	2.19	2.19
σ_e	0.78	0.78	0.78	0.77	1.08
ρ (Var. fraction due to α_i)	0.89	0.89	0.89	0.89	0.80

Robust standard errors clustered at the lease level in parentheses
Trend omitted due to collinearity with “Months Producing” variable
FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 100 lags)
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 19 – *Oil Production Regression Results for Stripper Well Leases: Panel Fixed Effects Model Using 3-month Rolling Average Product and Price*

	All Leases **GLS, FE**	**All Leases** **FE, DK**	>150 **Sales** **FE, DK**	>200 **Sales** **FE, DK**	>250 **Sales** **FE, DK**
Log 3 Month Rolling Avg Price, ϵ_s	0.501***	0.501***	0.190***	0.055	0.040$^+$
	(0.032)	(0.081)	(0.043)	(0.041)	(0.023)
Months Producing	−9.3e-03***	−9.3e-03***	−5.0e-03***	−3.3e-03***	−2.9e-03***
	(5.4e-04)	(8.1e-04)	(4.5e-04)	(6.2e-04)	(4.1e-04)
Months Producing Sqrd	3.0e-06***	3.0e-06***	7.1e-07**	−2.3e-07	−1.2e-07
	(6.0e-07)	(2.2e-07)	(2.6e-07)	(4.7e-07)	(4.3e-07)
Shut-In Status	−5.0***	−5.0***	−7.1***	−7.9***	−8.0***
	(0.13)	(0.44)	(0.30)	(0.46)	(1.2)
Constant	0.87***	0.87***	2.28***	2.82***	3.37***
	(0.12)	(0.19)	(0.10)	(0.08)	(0.11)
No. Obs.	1,192,331	1,192,331	864,853	676,153	369,744
No. Groups	4,731	4,731	3,416	2,657	1,450
F	612	57	600	215	864
R^2_w	0.089	0.089	0.19	0.29	0.40
σ_α	3.74	3.74	2.49	2.21	2.19
σ_e	3.08	3.08	2.15	1.56	0.82
ρ (Var. fraction due to α_i)	0.60	0.60	0.57	0.67	0.88

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 100 lags)

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 20 – *Oil Production Regression Results for Stripper Well Leases: Panel Fixed Effects Model Using 12-month Rolling Average Product and Price*

	All Leases **GLS, FE**	**All Leases** **FE, DK**	>150 **Sales** **FE, DK**	>200 **Sales** **FE, DK**	>250 **Sales** **FE, DK**
Log 12 Month Rolling Avg Price, ϵ_s	0.336***	0.336***	0.089**	0.010	0.019
	(0.035)	(0.058)	(0.032)	(0.028)	(0.029)
Months Producing	−6.9e-03***	−6.9e-03***	−4.1e-03***	−3.4e-03***	−2.8e-03***
	(4.9e-04)	(6.3e-04)	(4.9e-04)	(5.7e-04)	(3.7e-04)
Months Producing Sqrd	2.2e-06***	2.2e-06***	6.3e-07	2.5e-07	−1.4e-07
	(5.4e-07)	(1.7e-07)	(4.2e-07)	(5.1e-07)	(4.0e-07)
Shut-In Status	−5.0***	−5.0***	−6.9***	−7.2***	−5.9***
	(0.13)	(0.81)	(0.69)	(0.89)	(1.3)
Constant	1.53***	1.53***	2.61***	3.06***	3.43***
	(0.12)	(0.13)	(0.09)	(0.10)	(0.12)
No. Obs.	1,152,006	1,152,006	835,319	652,745	356,700
No. Groups	4,731	4,731	3,416	2,657	1,450
F	536	81	190	390	894
R^2_w	0.14	0.14	0.29	0.38	0.30
σ_α	2.92	2.92	2.27	2.14	2.19
σ_e	2.24	2.24	1.57	1.16	0.78
ρ (Var. fraction due to α_i)	0.63	0.63	0.68	0.77	0.89

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 100 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 21 – *Oil Production Regression Results for Stripper Well Leases: First–Difference Regressions Averaging Production for Months with no Sales Report*

	All Leases	<150 **Sales**	>150 **Sales**	>200 **Sales**	>250 **Sales**
	OLS, FD	**OLS, FD**	**OLS, FD**	**OLS, FD**	**OLS, FD**
$\Delta Ln(P_{it})$	0.059***	0.086***	0.049**	0.042+	0.059+
	(0.015)	(0.021)	(0.018)	(0.021)	(0.030)
$\Delta MonthsProducing$	.	.	.	.	.
	.	.	.	.	.
$\Delta MonthsProducingSqrd$	5.2e-07	6.1e-06***	−7.0e-07	−7.7e-07+	−2.1e-06***
	(4.0e-07)	(9.8e-07)	(4.5e-07)	(4.4e-07)	(5.0e-07)
$\Delta Shut-In$	−12.0***	−11.7***	−12.2***	−12.2***	−12.4***
	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)
Constant	−0.004***	−0.007***	−0.003***	−0.003***	−0.002***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
No. Obs.	1,196,799	325,988	868,129	678,754	371,199
No. Clusters	4,731	1,304	3,416	2,657	1,450
F	25,962	11,551	15,134	11,419	5,734
R^2	0.47	0.57	0.43	0.43	0.47

Robust standard errors clustered at the lease level in parentheses

$\Delta MonthsProducing$ is omitted due to collinearity with the Constant term which results from the inclusion of a time trend in the orignal model

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 22 – *Oil Production Regression Results for Stripper Well Leases: Dynamic Panel Fixed Effects Model*

	>200 **Sales** **ARDL(1,9)**	>250 **Sales** **ARDL(1,9)**	<150 **Sales** **ARDL(1,9)**
Error Correction, LR			
Log Price, ϵ_s	0.057^{**}	0.026	0.619^{***}
	(0.021)	(0.024)	(0.044)
SR			
$\hat{\phi}$	-0.395^{***}	-0.402^{***}	-0.289^{***}
	(0.008)	(0.015)	(0.009)
$\Delta Ln(P_{t-1})$	-0.135^{**}	-0.083^{+}	-0.901^{***}
	(0.042)	(0.048)	(0.074)
$\Delta Ln(P_{t-2})$	0.701^{***}	0.573^{***}	3.589^{***}
	(0.156)	(0.172)	(0.278)
$\Delta Ln(P_{t-3})$	-1.415^{***}	-1.312^{**}	-7.806^{***}
	(0.370)	(0.400)	(0.663)
$\Delta Ln(P_{t-4})$	1.767^{**}	1.951^{**}	10.835^{***}
	(0.589)	(0.629)	(1.044)
$\Delta Ln(P_{t-5})$	-1.259^{*}	-1.783^{**}	-9.985^{***}
	(0.642)	(0.681)	(1.110)
$\Delta Ln(P_{t-6})$	0.461	1.036^{*}	6.139^{***}
	(0.474)	(0.501)	(0.792)
$\Delta Ln(P_{t-7})$	-0.045	-0.389	-2.440^{***}
	(0.226)	(0.240)	(0.366)
$\Delta Ln(P_{t-8})$	-0.020	0.091	0.568^{***}
	(0.063)	(0.067)	(0.099)
$\Delta Ln(P_{t-9})$	4.8e-03	-1.0e-02	-5.9e-02***
	(7.9e-03)	(8.4e-03)	(1.2e-02)
Months Producing	-1.2e-03***	-1.1e-03***	-2.0e-03***
	(1.4e-04)	(1.8e-04)	(1.8e-04)
Months Producing Sqrd	-2.3e-07	-1.7e-07	7.7e-07**
	(1.4e-07)	(1.7e-07)	(2.4e-07)
Shut-In Status	-6.58^{***}	-7.61^{***}	-4.20^{***}
	(0.12)	(0.21)	(0.11)
Constant	2.50^{***}	2.73^{***}	0.74^{***}
	(0.07)	(0.11)	(0.05)
σ_e	1.042	1.031	0.889

Robust standard errors clustered at the lease level in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 23 – *Results from Cox Proportional Hazard Analysis of Lease Survival Based on Age Since First Production (Jan. 1996 - May 2011)*

	Coefficient	**Hazard Ratio**
SW Indicator	-0.158***	0.854
	(0.022)	
Price	-0.175***	0.839
	(0.006)	
Price^3	5.33e-05***	1.000
	(4.09e-06)	
Price^4	-4.07e-07***	1.000
	(4.05e-08)	
Prod. Prior to Shut–In	-0.0113***	0.989
	(0.0012)	
Price×Production	1.11e-04***	1.045
	(1.114e-05)	
Time Varying Covariates		
Price^3	1.46e-06***	1.000
	(5.97e-07)	
Price^4	-1.29e-08***	1.000
	(6.42e-09)	
Leases	15,027	
No. of Shut–in Events	9,541	
Wald χ^2	4,556***	

Robust standard errors clustered at the lease level in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Time varying covariates were interacted with $ln(t)$.

Appendix F: Non–Stripper Oil Well Lease Summary Statistics and Regression Results

Table 24 – *Summary Statistics for Non–Stripper Well Leases with more than 150 Reported Sales (Jan. 1996 – May 2011)*

Variable	Mean	Std. Dev.	N
Avg Daily Prod, bbls/d	34.06	275.3	767,524
Age in May 2011	374	206	4,116
Shut-In Status	0.12	0.33	767,524
Price Net of Royalty	38.73	20.71	756,986
Royalty Rate	0.123	0.020	767,524
Total Quarterly Production	3,119	2,5088	253,184
Avg Quarterly Oil Price	38.20	20.01	249,708
Total Yearly Production	12,505	10,0260	62,376
Avg Yearly Oil Price	37.68	18.64	61,515
Number of Leases		4,116	

Table 25 – *Oil Production Regression Results for the Non–Stripper Well Lease Full Sample: OLS, BE, and RE Model Comparisons Treating no Sales Report for a Month as Zero Lease Production*

	OLS	**BE**	**RE**
Log Price, ϵ_s	0.044	−0.492***	−0.068***
	(0.047)	(0.149)	(0.017)
Months Producing	−8.2e-03***	−1.1e-02***	−7.7e-03***
	(5.3e-04)	(7.3e-04)	(4.8e-04)
Months Producing Sqrd	9.3e-06***	1.2e-05***	9.1e-06***
	(7.1e-07)	(9.3e-07)	(6.4e-07)
Shut-In Status	−7.815***	−11.019***	−3.382***
	(0.047)	(0.147)	(0.067)
Trend	0.008***	0.012***	0.003***
	(0.001)	(0.002)	(0.000)
Constant	−2.67***	−0.25	−2.35***
	(0.16)	(0.50)	(0.09)
No. Obs.	1,905,530	1,905,530	1,905,530
No. Groups		11,729	11,729
F	5,923	1,174	
R^2	0.18	0.33	
R^2_o		0.18	0.17
R^2_b		0.33	0.31
R^2_w		0.045	0.046
σ_α			3.94
σ_e			3.41
ρ (Var. fraction due to α_i)			0.57

Robust standard errors clustered at the lease level in parentheses

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 26 – *Oil Production Regression Results for the Non–Stripper Well Lease Full Sample: Panel Fixed Effects Model Treating no Sales Report for a Month as Zero Lease Production*

	GLS, FE	**FE, DK**	**FE–REGAR**
Log Price, ϵ_s	-0.070^{***}	-0.070	-0.081^{***}
	(0.017)	(0.105)	(0.011)
Months Producing	$-5.2\text{e-}03^{***}$	$-5.2\text{e-}03^{**}$	$-3.5\text{e-}03^{***}$
	(4.9e-04)	(1.8e-03)	(1.6e-04)
Months Producing Sqrd	$9.1\text{e-}06^{***}$	$9.1\text{e-}06^{***}$	$8.0\text{e-}06^{***}$
	(6.5e-07)	(7.2e-07)	(1.7e-07)
Shut-In Status	-3.326^{***}	-3.326^{***}	-3.840^{***}
	(0.068)	(0.219)	(0.017)
Constant	-2.86^{***}	-2.86^{***}	-3.12^{***}
	(0.09)	(0.10)	(0.02)
No. Obs.	1,905,530	1,905,530	1,893,801
No. Groups	11,729	11,729	11,727
F	691	326	15,665
R^2_w	0.046	0.046	0.032
σ_α	4.43	4.43	
σ_e	3.41	3.41	
ρ (Var. fraction due to α_i)	0.63	0.63	

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with “Months Producing” variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 75 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 27 – *Oil Production Regression Results for the Non–Stripper Well Lease Full Sample: Panel Fixed Effects Model Using Production Averaging to Account for Intermittent Sales of Cumulative Production*

	GLS, FE	**FE, DK**	**FE, DK**
Log Price, ϵ_s	-0.017^{*}	-0.017	
	(0.009)	(0.059)	
Log 2-Month Lag Price, ϵ_s			-0.006
			(0.069)
Months Producing	$-4.3\text{e-}03^{***}$	$-4.3\text{e-}03^{***}$	$-4.3\text{e-}03^{***}$
	(2.6e-04)	(7.0e-04)	(6.8e-04)
Months Producing Sqrd	$5.6\text{e-}06^{***}$	$5.6\text{e-}06^{***}$	$5.6\text{e-}06^{***}$
	(3.5e-07)	(2.0e-07)	(2.1e-07)
Shut-In Status	-12.152^{***}	-12.152^{***}	-12.135^{***}
	(0.045)	(0.495)	(0.493)
Constant	3.47^{***}	3.47^{***}	3.43^{***}
	(0.05)	(0.10)	(0.11)
No. Obs.	1,905,530	1,905,530	1,888,174
No. Groups	11,729	11,729	11,729
F	19,713	1,017	1,301
R^2_w	0.81	0.81	0.81
σ_α	2.55	2.55	2.55
σ_e	1.28	1.28	1.28
ρ (Var. fraction due to α_i)	0.80	0.80	0.80

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 75 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 28 – *Oil Production Regression Results for the Non–Stripper Well Lease Full Sample: Panel Fixed Effects Model Using Quarterly Aggregation*

	GLS, FE	**FE, DK**	**FE, DK**
Log Avg Quarterly Price, ϵ_s	0.043*	0.043	
	(0.019)	(0.129)	
Log 2-Quarter Lag Price, ϵ_s			−0.065
			(0.120)
Months Producing	−6.5e-03***	−6.5e-03***	−5.3e-03**
	(4.8e-04)	(1.9e-03)	(1.8e-03)
Months Producing Sqrd	1.1e-05***	1.1e-05***	1.0e-05***
	(6.6e-07)	(8.9e-07)	(8.5e-07)
Shut-In Status	−3.605***	−3.605***	−3.531***
	(0.078)	(0.305)	(0.281)
Constant	2.49***	2.49***	2.54***
	(0.09)	(0.17)	(0.16)
No. Obs.	628,411	628,411	610,669
No. Groups	11,724	11,724	11,724
F	610	397	1,515
R^2_w	0.062	0.062	0.059
σ_α	4.13	4.13	4.17
σ_e	3.07	3.07	3.05
ρ (Var. fraction due to α_i)	0.64	0.64	0.65

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with “Months Producing” variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 20 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 29 – *Oil Production Regression Results for the Non–Stripper Well Lease Full Sample: Panel Fixed Effects Model Using Yearly Aggregation*

	GLS, FE	**FE, DK**	**FE, DK**
Log Avg Yearly Price, ϵ_s	-0.010	-0.010	
	(0.028)	(0.160)	
Log Avg 2-Year Price, ϵ_s			0.025
			(0.124)
Months Producing	$-8.6\text{e-}03^{***}$	$-8.6\text{e-}03^{***}$	$-8.0\text{e-}03^{***}$
	(5.1e-04)	(1.9e-03)	(1.5e-03)
Months Producing Sqrd	$1.3\text{e-}05^{***}$	$1.3\text{e-}05^{***}$	$1.2\text{e-}05^{***}$
	(6.7e-07)	(8.2e-07)	(4.3e-07)
Shut-In Status	-3.354^{***}	-3.354^{***}	-2.883^{***}
	(0.084)	(0.562)	(0.480)
Constant	5.26^{***}	5.26^{***}	5.03^{***}
	(0.09)	(0.10)	(0.09)
No. Obs.	155,284	155,284	136,838
No. Groups	11,698	11,698	11,698
F	512	1,903	2,619
R_w^2	0.082	0.082	0.061
σ_α	3.62	3.62	3.72
σ_e	2.67	2.67	2.59
ρ (Var. fraction due to α_i)	0.65	0.65	0.67

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with “Months Producing” variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 10 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 30 – *Oil Production Regression Results for Non–Stripper Well Leases With More Than 100 Reported Sales: Panel Fixed Effects Model, Averaging Production for Months with no Sales Report*

	GLS, FE	**FE, DK**	**FE, DK**	**FE, DK**	**FE, DK**
Log Price, ϵ_s	-0.036^{***}	-0.036^{+}			
	(0.009)	(0.020)			
Log 2-Month Lag Price, ϵ_s			-0.021		
			(0.025)		
Log 4-Month Lag Price, ϵ_s				-0.026	
				(0.034)	
Log Avg 12 Month Lag Price, ϵ_s					-0.154^{+}
					(0.082)
Months Producing	$-4.2\text{e-}03^{***}$	$-4.2\text{e-}03^{***}$	$-4.3\text{e-}03^{***}$	$-4.3\text{e-}03^{***}$	$-5.3\text{e-}03^{***}$
	(2.8e-04)	(4.7e-04)	(3.7e-04)	(3.4e-04)	(8.5e-04)
Months Producing Sqrd	$3.4\text{e-}06^{***}$	$3.4\text{e-}06^{***}$	$3.4\text{e-}06^{***}$	$3.3\text{e-}06^{***}$	$3.7\text{e-}06^{**}$
	(3.6e-07)	(3.7e-07)	(3.7e-07)	(3.7e-07)	(1.1e-06)
Shut-In Status	-13.26^{***}	-13.26^{***}	-13.23^{***}	-13.21^{***}	-5.01^{***}
	(0.068)	(0.354)	(0.352)	(0.351)	(0.653)
Constant	4.72^{***}	4.72^{***}	4.69^{***}	4.69^{***}	0.58^{**}
	(0.05)	(0.06)	(0.08)	(0.09)	(0.21)
No. Obs.	1,179,601	1,179,601	1,168,268	1,156,727	1,109,620
No. Groups	6,489	6,489	6,489	6,489	6,489
F	10,406	5,047	2,560	2,454	593
R^2_w	0.83	0.83	0.83	0.83	0.10
σ_α	2.38	2.38	2.38	2.39	3.45
σ_e	1.12	1.12	1.12	1.11	2.93
ρ (Var. fraction due to α_i)	0.82	0.82	0.82	0.82	0.58

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 75 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 31 – *Oil Production Regression Results for Non–Stripper Well Leases With Less Than 100 Reported Sales: Panel Fixed Effects Model, Averaging Production for Months with no Sales Report*

	GLS, FE	**FE, DK**	**FE, DK**	**FE, DK**	**FE, DK**
Log Price, ϵ_s	0.034*	0.034			
	(0.017)	(0.117)			
Log 2-Month Lag Price, ϵ_s			0.036		
			(0.126)		
Log 4-Month Lag Price, ϵ_s				0.023	
				(0.126)	
Log Avg 12 Month Lag Price, ϵ_s					0.014
					(0.158)
Months Producing	-2.5e-03***	-2.5e-03	-2.3e-03	-2.0e-03	-1.5e-03
	(5.3e-04)	(1.9e-03)	(1.9e-03)	(1.9e-03)	(2.1e-03)
Months Producing Sqrd	6.9e-06***	6.9e-06***	7.0e-06***	7.1e-06***	7.4e-06***
	(7.2e-07)	(9.8e-07)	(9.7e-07)	(9.8e-07)	(1.0e-06)
Shut-In Status	-11.4***	-11.4***	-11.4***	-11.4***	-11.4***
	(0.051)	(0.175)	(0.179)	(0.181)	(0.189)
Constant	1.26***	1.26***	1.19***	1.16***	1.02***
	(0.10)	(0.21)	(0.18)	(0.16)	(0.13)
No. Obs.	718,057	718,057	712,106	706,048	681,445
No. Groups	5,194	5,194	5,194	5,194	5,194
F	15,586	3,408	3,865	4,258	4,462
R^2_w	0.80	0.80	0.80	0.80	0.80
σ_α	2.56	2.56	2.59	2.62	2.71
σ_e	1.45	1.45	1.45	1.45	1.44
ρ (Var. fraction due to α_i)	0.76	0.76	0.76	0.77	0.78

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 75 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 32 – *Oil Production Regression Results for Non–Stripper Well Leases With More Than 150 Reported Sales: Panel Fixed Effects Model, Averaging Production for Months with no Sales Report*

	GLS, FE	**FE, DK**	**FE, DK**	**FE, DK**	**FE, DK**
Log Price, ϵ_s	-0.038^{***}	-0.038^{**}			
	(0.011)	(0.014)			
Log 2-Month Lag Price, ϵ_s			-0.017		
			(0.018)		
Log 4-Month Lag Price, ϵ_s				-0.019	
				(0.023)	
Log Avg 12 Month Lag Price, ϵ_s					-0.124^{***}
					(0.018)
Months Producing	$-3.9\text{e-}03^{***}$	$-3.9\text{e-}03^{***}$	$-4.0\text{e-}03^{***}$	$-4.0\text{e-}03^{***}$	$-4.2\text{e-}03^{***}$
	(3.0e-04)	(3.8e-04)	(3.2e-04)	(2.6e-04)	(1.8e-04)
Months Producing Sqrd	$2.7\text{e-}06^{***}$	$2.7\text{e-}06^{***}$	$2.7\text{e-}06^{***}$	$2.6\text{e-}06^{***}$	$2.1\text{e-}06^{***}$
	(3.5e-07)	(3.1e-07)	(3.2e-07)	(3.2e-07)	(2.6e-07)
Shut-In Status	-14.18^{***}	-14.18^{***}	-14.16^{***}	-14.13^{***}	-7.91^{***}
	(0.096)	(0.489)	(0.493)	(0.499)	(0.612)
Constant	5.33^{***}	5.33^{***}	5.30^{***}	5.29^{***}	1.92^{***}
	(0.05)	(0.04)	(0.05)	(0.05)	(0.07)
No. Obs.	756,986	756,986	749,451	741,802	710,584
No. Groups	4,116	4,116	4,116	4,116	4,116
F	6,033	2,694	2,887	3,273	985
R^2_w	0.84	0.84	0.84	0.83	0.25
σ_α	2.21	2.21	2.22	2.22	2.55
σ_e	0.95	0.95	0.95	0.94	2.05
ρ (Var. fraction due to α_i)	0.84	0.84	0.85	0.85	0.61

Robust standard errors clustered at the lease level in parentheses
Trend omitted due to collinearity with "Months Producing" variable
FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 75 lags)
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 33 – *Oil Production Regression Results for Non–Stripper Well Leases With More Than 180 Reported Sales: Panel Fixed Effects Model, Averaging Production for Months with no Sales Report*

	GLS, FE	**FE, DK**	**FE, DK**	**FE, DK**	**FE, DK**
Log Price, ϵ_s	-0.050^{***}	-0.050^{*}			
	(0.013)	(0.021)			
Log 2-Month Lag Price, ϵ_s			-0.030		
			(0.026)		
Log 4-Month Lag Price, ϵ_s				-0.027	
				(0.029)	
Log Avg 12 Month Lag Price, ϵ_s					-0.036
					(0.033)
Months Producing	$-3.1\text{e-}03^{***}$	$-3.1\text{e-}03^{***}$	$-3.3\text{e-}03^{***}$	$-3.3\text{e-}03^{***}$	$-3.9\text{e-}03^{***}$
	(3.5e-04)	(2.5e-04)	(2.3e-04)	(1.9e-04)	(2.2e-04)
Months Producing Sqrd	$2.2\text{e-}06^{***}$	$2.2\text{e-}06^{***}$	$2.2\text{e-}06^{***}$	$2.2\text{e-}06^{***}$	$2.2\text{e-}06^{***}$
	(3.9e-07)	(3.0e-07)	(3.0e-07)	(2.9e-07)	(1.1e-07)
Shut-In Status	-14.83^{***}	-14.83^{***}	-14.78^{***}	-14.75^{***}	-10.87^{***}
	(0.194)	(0.781)	(0.785)	(0.791)	(0.203)
Constant	5.66^{***}	5.66^{***}	5.64^{***}	5.63^{***}	2.41^{***}
	(0.07)	(0.06)	(0.06)	(0.06)	(0.07)
No. Obs.	449,866	449,866	445,043	440,192	420,736
No. Groups	2,432	2,432	2,432	2,432	2,432
F	1,566	7,656	7,987	7,824	1,889
R^2_w	0.76	0.76	0.75	0.75	0.45
σ_α	2.14	2.14	2.14	2.14	2.15
σ_e	0.82	0.82	0.82	0.81	1.14
ρ (Var. fraction due to α_i)	0.87	0.87	0.87	0.87	0.78

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 75 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 34 – *Oil Production Regression Results for Non–Stripper Well Leases: Panel Fixed Effects Model Using 3-month Rolling Average Product and Price*

	All Leases **GLS, FE**	**All Leases** **FE, DK**	> 100 **Sales** **FE, DK**	> 150 **Sales** **FE, DK**	> 180 **Sales** **FE, DK**
Log 3 Month Rolling Avg Price, ϵ_s	0.046^{*}	0.046	-0.080^{*}	-0.073^{***}	-0.060^{**}
	(0.018)	(0.122)	(0.039)	(0.011)	(0.021)
Months Producing	$-6.6\text{e-}03^{***}$	$-6.6\text{e-}03^{***}$	$-5.5\text{e-}03^{***}$	$-4.5\text{e-}03^{***}$	$-3.7\text{e-}03^{***}$
	(4.8e-04)	(1.7e-03)	(5.7e-04)	(4.2e-04)	(4.5e-04)
Months Producing Sqrd	$1.1\text{e-}05^{***}$	$1.1\text{e-}05^{***}$	$5.9\text{e-}06^{***}$	$3.7\text{e-}06^{***}$	$2.9\text{e-}06^{***}$
	(6.6e-07)	(7.5e-07)	(1.3e-06)	(4.7e-07)	(5.0e-07)
Shut-In Status	-3.4^{***}	-3.4^{***}	-5.4^{***}	-7.6^{***}	-8.9^{***}
	(0.072)	(0.38)	(0.20)	(0.42)	(1.4)
Constant	-2.04^{***}	-2.04^{***}	0.84^{***}	1.90^{***}	2.38^{***}
	(0.09)	(0.14)	(0.15)	(0.04)	(0.07)
No. Obs.	1,888,174	1,888,174	1,168,268	749,451	445,043
No. Groups	11,729	11,729	6,489	4,116	2,432
F	638	1,095	278	1,507	9,603
R^2_w	0.060	0.060	0.19	0.41	0.48
σ_α	4.15	4.15	3.10	2.45	2.17
σ_e	3.05	3.05	2.18	1.44	0.91
ρ (Var. fraction due to α_i)	0.65	0.65	0.67	0.74	0.85

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 75 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 35 – *Oil Production Regression Results for Non–Stripper Well Leases: Fixed Effects Model Using 12-month Rolling Average Product and Price*

	All Leases GLS, FE	**All Leases FE, DK**	>100 **Sales FE, DK**	>150 **Sales FE, DK**	>180 **Sales FE, DK**
Log 12 Month Rolling Avg Price, ϵ_s	-0.059^{*}	-0.059	-0.190^{**}	-0.144^{*}	-0.108^{+}
	(0.025)	(0.139)	(0.073)	(0.065)	(0.061)
Months Producing	$-8.1\text{e-}03^{***}$	$-8.1\text{e-}03^{***}$	$-7.4\text{e-}03^{***}$	$-6.7\text{e-}03^{***}$	$-4.1\text{e-}03^{***}$
	(5.1e-04)	(1.7e-03)	(1.1e-03)	(9.2e-04)	(4.5e-04)
Months Producing Sqrd	$1.3\text{e-}05^{***}$	$1.3\text{e-}05^{***}$	$8.7\text{e-}06^{***}$	$6.4\text{e-}06^{***}$	$3.8\text{e-}06^{***}$
	(6.8e-07)	(7.9e-07)	(1.3e-06)	(6.1e-07)	(4.1e-07)
Shut-In Status	-3.1^{***}	-3.1^{***}	-5.3^{***}	-7.1^{***}	-7.5^{***}
	(0.70)	(0.59)	(0.38)	(0.72)	(1.4)
Constant	-0.64^{***}	-0.64^{***}	1.79^{***}	2.56^{***}	2.53^{***}
	(0.09)	(0.11)	(0.10)	(0.14)	(0.09)
No. Obs.	1,807,607	1,807,607	1,115,586	714,536	423,168
No. Groups	11,729	11,729	6,489	4,116	4,432
F	493	820	1,097	629	2,898
R^2_w	0.073	0.073	0.26	0.43	0.39
σ_α	3.68	3.68	2.93	2.47	2.23
σ_e	2.57	2.57	1.74	1.26	0.91
ρ (Var. fraction due to α_i)	0.67	0.67	0.74	0.79	0.86

Robust standard errors clustered at the lease level in parentheses

Trend omitted due to collinearity with "Months Producing" variable

FE, DK: Panel Fixed Effects with Driscoll-Kraay standard errors (autocorrelation up to 75 lags)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 36 – *Oil Production Regression Results for Non–Stripper Well Leases: First–Difference Regressions Averaging Production for Months with no Sales Report*

	All Leases	<100 **Sales**	>100 **Sales**	>150 **Sales**	>180 **Sales**
	OLS, FD	**OLS, FD**	**OLS, FD**	**OLS, FD**	**OLS, FD**
$\Delta LN(P_{it})$	−0.121***	−0.127***	−0.119***	−0.108***	−0.065***
	(0.012)	(0.024)	(0.014)	(0.016)	(0.017)
$\Delta MonthsProducing$	.	.	.	.	.
	.	.	.	.	.
$\Delta MonthsProducingSqrd$	−4.7e-05***	−9.2e-05***	−1.8e-05***	−9.3e-06***	2.5e-06***
	(1.0e-06)	(2.2e-06)	(9.0e-07)	(9.4e-07)	(6.7e-07)
$\Delta Shut-In$	−11.4***	−10.9***	−12.0***	−12.4***	−12.3***
	(0.031)	(0.040)	(0.048)	(0.062)	(0.093)
Constant	0.042***	0.085***	0.015***	0.008***	−0.004***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
No. Obs.	1,896,914	715,100	1,173,978	753,245	447,462
No. Clusters	11,729	5,194	6,489	4,116	2,432
F	43,867	24,985	20,474	13,770	5,878
R^2	0.42	0.44	0.41	0.45	0.47

Robust standard errors clustered at the lease level in parentheses

$\Delta MonthsProducing$ is omitted due to collinearity with the Constant term which results from the inclusion of a time trend in the orignal model

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 37 – *Oil Production Regression Results for Non–Stripper Well Leases: Dynamic Panel Fixed Effects Model*

	>150 **Sales**	>180 **Sales**	<100 **Sales**
	ARDL(1,9)	**ARDL(1,9)**	**ARDL(1,9)**
Error Correction, LR			
Log Price, ϵ_s	-0.014	-0.006	0.058^{*}
	(0.018)	(0.023)	(0.029)
SR			
$\hat{\phi}$	-0.480^{***}	-0.439^{***}	-0.433^{***}
	(0.007)	(0.010)	(0.006)
$\Delta Ln(P_{t-1})$	-0.199^{***}	-0.189^{***}	-0.548^{***}
	(0.044)	(0.046)	(0.068)
$\Delta Ln(P_{t-2})$	0.531^{***}	0.528^{***}	2.063^{***}
	(0.153)	(0.159)	(0.256)
$\Delta Ln(P_{t-3})$	-1.075^{**}	-0.906^{*}	-4.923^{***}
	(0.352)	(0.372)	(0.612)
$\Delta Ln(P_{t-4})$	1.215^{*}	0.780	7.232^{***}
	(0.546)	(0.586)	(0.963)
$\Delta Ln(P_{t-5})$	-0.621	0.011	-6.828^{***}
	(0.579)	(0.633)	(1.020)
$\Delta Ln(P_{t-6})$	-0.051	-0.634	4.199^{***}
	(0.418)	(0.464)	(0.722)
$\Delta Ln(P_{t-7})$	0.218	0.552^{*}	-1.652^{***}
	(0.197)	(0.221)	(0.330)
$\Delta Ln(P_{t-8})$	-0.098^{+}	-0.205^{***}	0.382^{***}
	(0.055)	(0.061)	(0.088)
$\Delta Ln(P_{t-9})$	1.5e-02*	3.0e-02***	$-$4.0e-02***
	(6.8e-03)	(7.5e-03)	(1.0e-02)
Months Producing	$-$1.4e-03***	$-$1.2e-03***	$-$1.1e-03***
	(1.7e-04)	(1.8e-04)	(2.5e-04)
Months Producing Sqrd	6.2e-07***	5.3e-07**	4.0e-06***
	(1.8e-07)	(1.8e-07)	(3.3e-07)
Shut-In Status	-7.39^{***}	-7.73^{***}	-5.49^{***}
	(0.11)	(0.16)	(0.07)
Constant	2.53^{***}	2.48^{***}	0.45^{***}
	(0.05)	(0.07)	(0.05)
σ_e	1.229	1.087	1.582

Robust standard errors clustered at the lease level in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3 Predicting CO_2 Emissions in Developing Countries

Abstract

Forecasts of carbon emissions are critical inputs to climate models which are used to predict the possible impacts of climate change. Unfortunately, most prediction models used to date have consistently underestimated carbon emissions. To predict emissions from developed countries, this paper uses the relationship between per capita GDP and carbon dioxide (CO_2) emissions, an environmental Kuznets curve (EKC) analysis. To make forecasts for developing economies, however, I propose an alternative specification in which the independent variable of interest is a measure of the country's socio-economic status (SES) based on available surveys of household characteristics and possessions. A measure of household characteristics such as SES should more closely correlate to the consumption-induced emission of carbon dioxide than a highly aggregated measure such as per capita GDP would. Per capita GDP will mask potential widespread income inequalities, whereas household level data avoids this problem. Using SES produces a statistically significant quadratic increase in emissions for countries as SES increases and improves on in-sample prediction of carbon emissions which rely on GDP per capita alone.

3.1 Introduction

Forecasting CO_2 emissions has been the focus of many prominent studies (Nordhaus and Yohe 1983; Reilly et al. 1987; IPCC 1990; Manne and Richels 1992; IPCC 2001). These forecasts rely on models of per capita GDP growth rates which attempt to take into account changes in those sectors that contribute to carbon emissions using various assumptions about technological change and population growth. While the models vary in methodology and complexity, they consistently underestimate global carbon emissions. The IPCC (2007) notes that there are nearly 400 emission forecasts (they refer to the them as scenarios) assuming no global policy intervention. Figures 11 and 12 compare the primary scenarios to the actual level of emissions. Notably, the predictions by Holtz-Eakin and Selden (1995) based on a much less computationally intensive method appear to bracket the actual level of emissions (referred to as HE&S Level and HE&S Log). Their predictions use a reduced form model of per capita GDP and per capita carbon dioxide emissions of the type used in environmental Kuznets curve analyses.

The relationship between per capita income and environmental pollutants is based on Simon Kuznets' work on income inequality (1955). This type of analysis has generated a prolific amount of empirical research in the past twenty years. Researchers hypothesize that as a country initially develops, pollutant emissions will increase along with economic development until the costs of the resulting environmental degradation, to both health and environmental quality, are taken into account by a wealthier population. The demand for improved health and less pollution thereby increases, causing per capita emissions to peak and eventually decrease. This is known as the environmental Kuznets curve (EKC). This type of relationship between economic development, as measured by real GDP per capita, and per capita emissions has been verified for a wide range of pollutants using reduced-form models (e.g. Grossman and Krueger 1995; Selden and Song 1994; List and Gallet 1999; Millimet et al. 2003). A survey of much of this literature is found in Panayotou (2000) as well as Yavapolkul (2005). For pollutants such as CO_2 with long-term, global impacts, however, a monotonically increasing relationship is expected and has been confirmed by most studies (Holtz-Eakin and Selden 1995; Cole et al. 1997; Moomaw and Unruh 1997; Galeotti and Lanza 1999; Sun 1999; Friedl and Getzner 2003; Aldy 2005; Bertinelli and Strobl 2005; Dijkgraaf and Vollebergh 2005; Frankel and Rose 2005; Azomahou et al. 2006).

Holtz-Eakin and Selden (1995) use this EKC approach to estimate what per capita CO_2 emissions will be at various levels of per capita GDP and, combined with United Nations population projections and a simple forecast model of per capita GDP growth, they generate emissions paths which have largely bracketed actual emissions since 1985. Despite their relative "success" in predicting the range of CO_2 emissions over the short run compared to the more structural models shown, the rapid increase in emissions since 2003 has pushed total emissions outside of this range in 2007 (the most recent year for which data has been reported). Figure 13 shows the total level of carbon emissions, as well as the amount of emissions attributable to both developing and developed (OECD) countries. It is clear that while total emissions from developed countries have remained relatively flat since the early 1990s, the total emissions from developing countries have increased rapidly, particularly since 2000.

In an attempt to improve on these forecasts for developing countries, I propose a modification to Holtz-Eakin and Selden's EKC approach. Per capita GDP is a highly aggregated measure which may not have as much meaning for developing countries. It can mask differences between countries with households that have a fairly consistent level of wealth and those where resource wealth is highly concentrated while large fractions of the population live at the subsistence level. In this latter group of countries, emissions incurred by consumption are likely to be very low. This is evident from an inspection of the unconditional emission vs. GDP relationship in Figure 14, where the same level of per capita GDP can be associated with a wide range of per capita emissions. To overcome this problem, some measure of household level data should correspond more closely to emissions. Since household income and expenditure data are difficult to measure and interpret given the problems laid out in Vyas and Kumaranayake (2006) and Gwatkin et al. (2007), data on observable household characteristics and possessions has been extensively used since the late 1990s. For the subsample of developing economies I use this type of household data to create country socio-economic status (SES) indicators for those years with available data. I then use this SES measure in a Kuznets curve model instead of per capita GDP.

Using this SES measure produces a statistically significant quadratic increase in CO_2 emissions for developing countries as SES increases. Further, by incorporating this SES–emissions relationship for developing countries into predictions of global emissions, in-sample predictions appear to be more accurate than just relying on GDP per capita alone, although forecasting assumptions given data limitations would render tests of this to have relatively low power.

This paper is organized as follows. Section 3.2 discusses the previous literature aimed at forecasting global carbon emissions. Section 3.3 discusses the empirical analysis used in the present study. Sections 3.4 and 3.5 describe the data and results, respectively. Section 3.6 concludes.

3.2 Literature Review

Several studies have undertaken the task of forecasting future emissions of CO_2 (Nordhaus and Yohe 1983; Reilly et al. 1987; IPCC 1990; Manne and Richels 1992; IPCC 2001). The Intergovernmental Panel on Climate Change (IPCC) has been the primary source of emissions forecasts since 1992 when they included six scenarios of future emissions in an assessment report (IPCC 1992). Currently, the IPCC uses a set of scenarios developed to capture a range of potential drivers of greenhouse gas (GHG) emissions (IPCC 2000). The IPCC groups the scenarios into four categories (A1, A2, B1, and B2). They describe these categories thusly:

> The A1 storyline assumes a world of very rapid economic growth, a global population that peaks in mid-century and rapid introduction of new and more efficient technologies. A1 is divided into three groups that describe alternative directions of technological change: fossil intensive (A1FI), non-fossil energy resources (A1T) and a balance across all sources (A1B). B1 describes a convergent world, with the same global population as A1, but with more rapid changes in economic structures toward a service and information economy. B2 describes a world with intermediate population and economic growth, emphasizing local solutions to economic, social, and environmental sustainability. A2 describes a very heterogeneous world with high population growth, slow economic development and slow technological change.

With the assumptions from these families of scenarios, the estimated emissions paths are then forecast using a number of models[24] of economic growth. These models have assumptions about GDP growth in various sectors of country or regional economies. Using the scenario assumptions on population growth, economic growth, and technological change, the models then forecast what emissions will be. The dashed lines in Figures 11 and 12

[24] The models include: (1) Asian Pacific Integrated Model (AIM), (2) Atmospheric Stabilization Framework Model (ASF), (3) Integrated Model to Assess the Greenhouse Effect (IMAGE), (4) Multiregional Approach for Resource and Industry Allocation (MARIA), (5) Model for Energy Supply Strategy Alternatives and their General Environmental Impact (MESSAGE), and (6)The Mini Climate Assessment Model (MiniCAM).

show the forecast emissions[25] from fossil fuel and industrial processes over the medium and short terms, respectively, from one scenario in each of the four categories. The $A1$ and $B1$ scenarios represent the best and worst cases, respectively, and fall well outside of the bulk of scenario forecasts. The $A2$ and $B2$ scenarios are more in line with the median IPCC forecasts. While the IPCC does not take a position on which of the scenarios represent the likely path of emissions, their approach in general seems to be to cast a huge net of possibilities, thus assuring that the actual lies somewhere within. In this instance, the actual does lie within the best and worst case scenarios, however, preliminary data from 2010 indicate that emissions have since exceeded the worst case.

Contrasted with the approach of the IPCC discussed above, Holtz-Eakin and Selden (1995) use an innovative approach that is much less reliant on assumptions about growth in different sectors of national economies, how each of those sectors will contribute to emissions, and how technology will evolve over time. Given a forecast of country populations, the approach taken by Holtz-Eakin and Selden (1995) relies on only two assumptions: (1) GDP growth rates will tend to converge, such that future growth will follow past growth trends, and (2) the underlying relationship between per capita GDP and per capita CO_2 will be the same across countries. The former is estimated with a simple model of GDP growth and the latter using an environmental Kuznets curve model. Their approach will be discussed in more detail in Section 3.3.1. Results of their analysis compared to the IPCC forecasts are displayed in Figures 11 and 12, along with earlier forecasts by Nordhaus and Yohe (1983), Reilly et al. (1987), and Manne and Richels (1992). The actual atmospheric concentration of CO_2 and IPCC forecast concentration are shown in Figure 11 for reference. The Holtz-Eakin and Selden forecasts from regressions in logs and levels largely bracket actual global emissions, however, did not pick up the large growth in emissions beginning in 2003. The IPCC worst case scenario does predict such an increase, but it is unclear from the published results if the IPCC forecast growth was due to increased emissions in developing countries, developed countries, or both. From Figure 13 it is clear that the increase in emissions in recent years is primarily from developing countries.

[25] Results are generally displayed as "Carbon Emissions" rather than "CO_2 Emissions". To convert the latter to the former, simply divide by 3.664 (the mass of carbon per unit of CO_2).

3.3 Empirical Analysis

3.3.1 Predicting Carbon Emissions from GDP Forecasts

Holtz-Eakin and Selden's method is computationally more attractive than those used in the IPCC forecasts because it relies primarily on past data to make its forecasts, rather than the unknowables in the IPCC approach. The first requirement for forecasting CO_2 emissions is to develop some relationship between the level of emissions and economic conditions in each county. An EKC analysis provides this type of a relationship. EKC theory suggests that as a country initially develops, pollutant emissions will increase along with economic development until the costs of the resulting environmental degradation, to both health and environmental quality, are taken into account by a wealthier population. The demand for improved health and less pollution thereby increases, causing per capita emissions to peak and eventually decrease. This hypothesis would lead to a relationship that is quadratic in per capita GDP. Some authors suggest that a cubic relationship could result if emissions begin to increase after an initial plateau as new technologies may be discovered that lead to increases in emissions through consumption of new goods. This new consumption may offset any increases in efficiencies the new technology may afford. Incorporating the potential for a quadratic or cubic relationship, the basic reduced form model is:

$$E_{it} = \beta_0 + \beta_1 Y_{it} + \beta_2 Y_{it}^2 + \beta_3 Y_{it}^3 + f_i + \gamma_t + \epsilon_{it}, \tag{9}$$

where E_{it} is the per capita environmental indicator, in this case metric tons (mt) of CO_2 emissions per person, of country i in year t, Y_{it} is the per capita GDP of country i in year t, f_i are fixed effects for country i, and γ_t is a linear time trend which accounts for factors such as technological change affecting all countries.

Holtz-Eakin and Selden (1995) point out that per capita GDP is used "to assess both the direct and indirect consequence of growth, [therefore] those variables that are endogenous consequences of growth–e.g. the composition of output, regulations and taxes influencing fossil fuel consumption, patterns of urbanization and sub-urbanization, etc.–should be omitted from this simple model." As a result, we are not able to attain causal linkages, but can understand the underlying relationship between growth and emissions, controlling for country-specific and time-specific factors.

The estimated model from equation (9) can be used to forecast future carbon emissions if forecasts of GDP can be made. Holtz-Eakin and Selden (1995, p. 92) forecast GDP growth for all countries using the estimated equation:

$$\widehat{ln(y_{it+1}) - ln}(y_{it}) = 0.0178 + 0.00822 ln(y_{it}) - 0.00212[ln(y_{it})]^2 + \tau_{it}. \tag{10}$$

While they do not define τ_{it}, it appears to represent a forecast error term. This model captures the theory of convergence in growth rates between developed and developing countries. Growth rates initially expand as a country develops, then plateau and begin to fall. I re-estimate their model using the data as described in section 3.4 and use it to make in-sample predictions of aggregate carbon emissions starting in 1991. Making these predictions is a simple matter of first using the per capita GDP for country i in 1990 in the re-estimated version of equation (10) to obtain a forecast of per capita GDP in 1991. With this forecast of GDP, I forecast per capita CO_2 using estimates of equation (9). To obtain each country's carbon emission I multiply this result by country i's population in 1991 and divide by 3.664 (the amount of carbon by mass in each unit of CO_2). Annual global emissions are then the aggregate of each country's carbon emissions calculated by this method for each year.

3.3.2 Predicting Carbon Emissions from a SES Measure for Developing Countries

Per capita GDP is a highly aggregated measure which may not have as much meaning for developing countries. It can mask differences between countries with households that have a fairly consistent level of wealth and those where resource wealth is highly concentrated while large fractions of the population live at the subsistence level. In this latter group of countries, emissions incurred by consumption are likely to be very low. This is evident from an inspection of the unconditional GDP–emissions relationship in Figure 14, where the same level of per capita GDP can be associated with a wide range of per capita emissions.

To overcome this problem, a measure of household level data that still captures the relative wealth of countries should correspond more closely to emissions. Like GDP, a measure of household level data will account for emissions from the manufacture and consumption of domestically produced goods and should partially account for emissions from the manufacture of exported goods through wages paid to households. The wages will be used to expand household consumption, which will be picked up by household level data. What will not be picked up by this data, however, are emissions from manufacturing where workers are paid only a subsistence wage that does not result in discernible changes to household characteristics or possessions. Since household income and expenditure data are difficult to measure and interpret given the problems laid out in Vyas and Kumaranayake (2006) and Gwatkin et al. (2007), data on observable household characteristics and possessions has been extensively used since the late 1990s to assess the wealth of those households.

To improve on the shortcomings of using the GDP–emissions relationship to estimate emissions in developing countries, I therefore use household characteristic and possession data to create country socio-economic status (SES) indicators. Vyas and Kumaranayake (2006), Houweling et al. (2003), and McKenzie (2003), for example, discuss how these types of indicators can be generated at the household level and used to assess outcomes across households. In this paper I generate SES measures at the country level and use them to assess environmental outcomes. To create this SES measure I combine data on household characteristics and possessions using weightings from a principal component analysis (PCA) as described in Vyas and Kumaranayake (2006). The components are created by applying a specific weighting to each of the household variables as follows:

$$
\begin{aligned}
PC_1 &= a_{11}X_1 + a_{12}X_2 + \ldots + a_{1n}X_n, \\
&\vdots \\
PC_n &= a_{n1}X_1 + a_{n2}X_2 + \ldots + a_{nn}X_n.
\end{aligned}
$$

$PC_1,\ldots,PC_n$ is the vector of principal components, $X_1,\ldots,X_n$ is the vector of household variables being combined, and $a_{11},\ldots,a_{nn}$ is the matrix of weights associated with each principal component and variable. In a PCA using unstandardized data, the weights are the eigenvectors of the data's correlation matrix. Each principal component then accounts for a portion of the total variation in the original data. The amount of variation associated with each of the n principal components is calculated as the eigenvalue, λ, of each eigenvector divided by n. The components are ordered so that PC_1 explains the largest variation and PC_n explains the smallest. PC_1 is then taken to represent the country's socio-economic status for the year when the data was collected.

To determine the relationship between socio-economic status and per capita CO_2 emissions, I replace Y_{it} in equation (9) with the SES measure of country i in year t:

$$E_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 S_{it}^2 + \beta_3 S_{it}^3 + \gamma_t + \epsilon_{it}. \tag{11}$$

Given the data limitations described in the next section, I am unable to use a fixed effects specification, however, this appears to be unnecessary.

Making in-sample predictions of CO_2 emissions using the results of equation (11) follows a method similar to that described in section 3.3.1. Unfortunately, the frequency of household surveys used to generate data for the SES analysis precludes a forecast of the type used for per capita GDP. Fortunately, however, the SES appears to grow in a roughly linear fashion in each country as described in section 3.5.2. I therefore use a simple linear growth path of the SES measure for each country to calculate in-sample carbon emissions predictions for the developing countries that have been surveyed, combined with the GDP analysis described in section 3.3.1 for all other countries.

3.4 Data

The per capita CO_2 emissions data were obtained from the Oak Ridge National Laboratory (ORNL) for 162 countries from 1980 through 2007. This is the most comprehensive source of CO_2 emissions data and is the primary source for previous studies. The data includes CO_2 emitted by fossil fuel consumption, the manufacture of cement, and gas flaring, but notably omits the difficult to quantify emissions from changing patterns of land use. The latter could be significant for developing countries and if good estimates of this could be obtained it would be of value for future studies. Per capita GDP were obtained for the same countries and years from Penn World Table Version 7.0. Missing GDP data were obtained from the World Bank's World Development Indicators. Yearly country populations were also obtained from the World Bank's World Development Indicators. The GDP data are expressed on a purchasing power parity basis in 2005 dollars to account for inflation and currency purchasing power. The result is a panel over 28 years consisting of 4,190 total observations.

Household characteristic and possession data used to generate the SES indicator were collected by the Demographic and Health Survey (DHS) program. This program is a collaboration between Macro International and the U.S. Agency for International Development. The surveys were conducted between 1990 and 2007 in 75 developing countries throughout the world. The survey contains data relating to the construction of the house (only information on flooring material was used in this study), the source of water, household sanitation facilities, whether or not the house has electricity, and the presence of possessions such as radios, telephones, televisions, bicycles, motorcycles, refrigerators, and automobiles. Table 38 provides a summary of the data on the 20 variables used to generate the SES indicator. Each number represents the percentage of households which have the characteristic or possession. Only the surveys from 63 countries contained all of the information required to calculate an indicator. The result is a pooled sample of 140 observations on the 20 variables. A complete listing of the 63 countries is included in Table 39 in Appendix G. The available data was collected at varying frequencies for the countries, so only one year is available for some, while five years are available for others. Further, the years when data was collected are not consistent across countries.

3.5 Results

3.5.1 Results of GDP EKC Analysis

I first generate pooled ordinary least squares (OLS) estimates and robust standard errors of equation (9). I test for and confirm serial correlation in the residuals and rerun the

OLS regression with robust standard errors clustered at the country level. I then estimate the model using a fixed effects GLS procedure with heteroskedastic errors and an AR(1) autocorrelation process, allowing for country-specific AR(1) parameters, ρ_i. Stern et al. (1996) noted that simultaneity may be a problem if environmental degradation affects economic output. Therefore, following Cole et al. (1997) I use a Hausman test to verify the exogenity of current GDP in the EKC model, where lagged income is the instrumental variable. OLS and panel GLS estimates are shown in Table 40. To highlight the problems associated with using GDP to predict emissions in developing countries, I re-estimate model (9) for those countries with less than $10,000 per capita GDP. Figure 14 compares the resulting predicted per capita CO_2 emissions from OLS and GLS fixed–effects estimations to the actual emissions in each country each year (the prediction uses the average country effect).

I next re-estimate equation (10) for GDP growth using a panel fixed effects regression. The results of this regression, using robust standard errors clustered at the country level, are presented in Table 42. The coefficients are similar in magnitude to those obtained by Holtz-Eakin and Selden (1995) with data between 1951 and 1986.[26] From this analysis, GDP growth rates are forecast to peak at approximately $9,300 (2005 US dollars) and fall slowly thereafter. Combined with estimates of equation (9) and country populations, I calculate the path of total carbon emissions between 1991 and 2008. The results are plotted in Figure 19 and labeled as "GDP (HE&S Method)."

3.5.2 Results of SES Analysis

Table 38 shows the weightings for the first principal component of each of the household characteristics and possessions used to generate the socio-economic status measure. Figure 15 plots the calculated SES indicator for each country and year versus per capita CO_2 emissions. A relatively smooth trend with few outliers is evident. Figure 16 compares the per capita emissions versus SES relationship to the emissions versus per capita GDP

[26] The coefficients are not directly comparable for two reasons. First, Holtz-Eakin and Selden's data were in 1986 US dollars whereas mine are in 2005 US dollars. Second, Holtz-Eakin and Selden run regressions using natural logs of data in thousands of dollars, whereas I use dollars. The constant and coefficient on $ln(y_{it})$ in Table 42 would therefore be expected to differ.

relationship for the same countries and years. While no unambiguous relationship between GDP and emissions is evident from the graph, a case could be made that emissions are tending to flatten out above roughly $3,000 per capita. Such is not the case using a household level measure of country wealth to predict emissions. Table 41 shows OLS estimates of equation (11), which is used to predict emissions depending on a country's SES. The resulting predicted emissions curve is shown in Figure 15.

In order to make predictions of global annual carbon emissions using this SES–CO_2 emissions relationship, a method must be devised to predict what each country's SES measure would have been between sample years, and also what it might be into the future. Figures 17 and 18 are plots of the SES measure over time for the 25 developing countries with survey data for three or more years. Trend lines are also displayed for each country, with the slope displayed in the legend. The plots are divided between non-African and African countries for ease of viewing. While the sample size is small and care must be taken in drawing conclusions, rough approximations of how the SES evolves over time can be made. Fifteen of these 25 countries display linear growth in SES with R^2 values greater than 0.90. Zambia shows a decline in SES over time. Of the remaining nine countries, only linear models of Jordan and Peru have R^2 less than 0.75. The average increase in the SES of African countries is approximately 0.074 per year (standard deviation of 0.04), while non-African countries experience a slightly higher 0.11 increase per year (standard deviation of 0.06). The overall average increase is 0.09 (standard deviation of 0.05). There are an additional 18 countries with only two data points each. The average rate of change in SES for these 18 countries is also a growth of 0.091 per year, however, there is a good deal of variation as the standard deviation for this group is 0.29.

The lack of data in this case suggests that the most appropriate SES growth forecast model is a simple one that can be refined after additional surveys are conducted. I therefore apply a 0.09 across the board growth in SES per year starting in 1991 for the 63 developing countries for which survey data is available. I then forecast carbon emissions for each of these countries each year from the estimates of equation (11). For all other countries, the results described in Section 3.5 are used. The global annual carbon emissions forecast using this method are also displayed in Figure 19. This method indicates an improvement over using per capita GDP to predict emissions in developing countries. A formal test of the difference between the two forecasts is not conducted given the very broad assumption about the growth in SES over time, which is likely to impart large errors in the resulting carbon emissions. Directionally, however, the closer approximation of actual emissions is not unexpected given the shape of the SES–emissions relationship compared to that of the GDP–emissions relationship in developing countries.

While the SES approach toward predicting carbon emissions did not forecast the rapid increase in emissions beginning in 2002 and 2003, there is a simple explanation for this: China. Household surveys in China have not been conducted as part of the DHS, which required using the GDP–emissions relationship instead. While it is unclear that the SES approach would have predicted the rapid increase even if survey data from China was available over several years before and after the increase, Figure 20 compares predictions using the two methods to actual emissions if China is ignored. Using the GDP–emissions relationship alone appears to yield good predictions up until 2002, at which point increasing emissions in the developing world, even without the contribution of China, cause it to under predict emissions. Using the SES–emissions relationship for developing countries appears to over predict emissions. This could be the result of an oversimplified forecast for SES growth.

3.6 Conclusions and Extensions

Rather than relying on complex models and guesses at possible futures to forecast carbon emissions, Holtz-Eakin and Selden (1995) use a simple computational approach with straightforward models of GDP growth and the relationship between income and emissions. I employ a similar approach, but attempt to improve on predictions of emissions in developing countries by using a measure of socio-economic status rather than GDP. This SES measure can be generated from household characteristics and possessions survey data using a principal component analysis, and to my knowledge has not been applied to the study of environmental outcomes previously.

Using this approach, I find that the SES–emissions relationship in developing countries is a relatively smooth, quadratically increasing function. This is in stark contrast to the ambiguous GDP–emissions relationship for the same countries. I also find that with an assumption of linear growth in a country's SES over time, this approach appears to improve in-sample predictions of global carbon emissions. While data limitations precludes formal testing of this proposition, these results provide a promising path for further study when additional surveys are conducted.

Further work in the forecast of carbon emissions using this method will require additional years of household survey data, as well as the inclusion of large emitters like China. Its absence from the DHS data is notable. This approach can potentially be applied to study the relationship between wealth and the emission of other pollutants. While we are not likely to see per capita carbon emissions peak and fall within a sample of developing countries, other pollutants with more local and directly discernible environmental and health impacts may behave in such a manner when compared to SES.

Appendix G: Figures

Figure 11 – *Comparing Prediction Models: Actual (1950 through 2008) and Predicted Global Carbon Emissions*

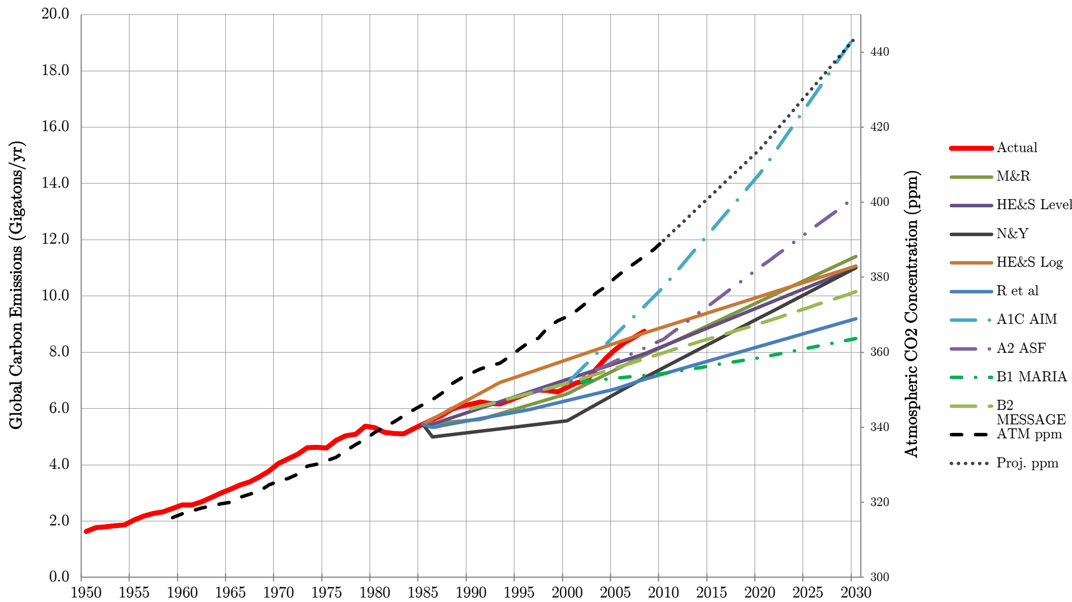


Figure 12 – *Comparing Prediction Models: Actual (1985 through 2008) and Predicted Global Carbon Emissions*

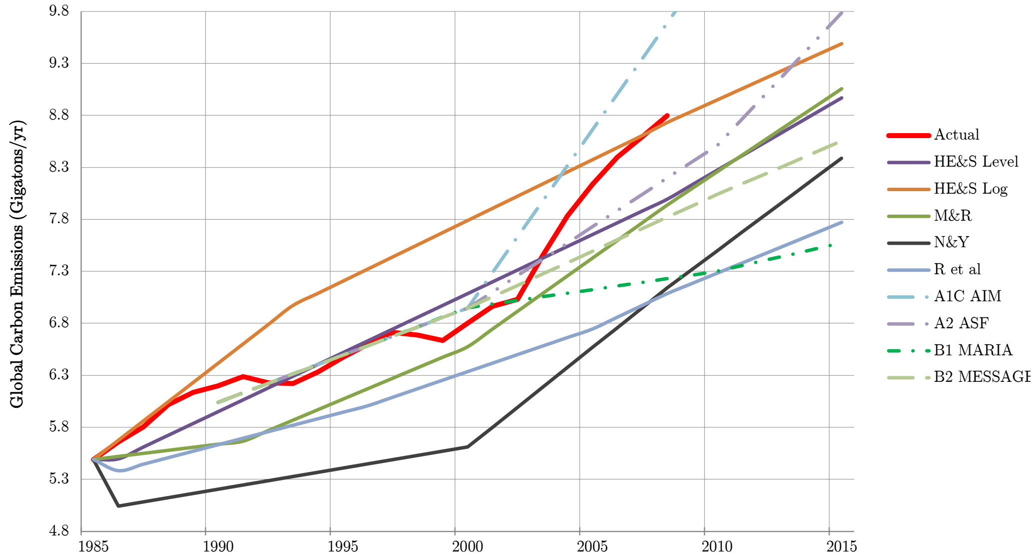


Figure 13 – *Global Carbon Emissions from Developing and Developed Countries*

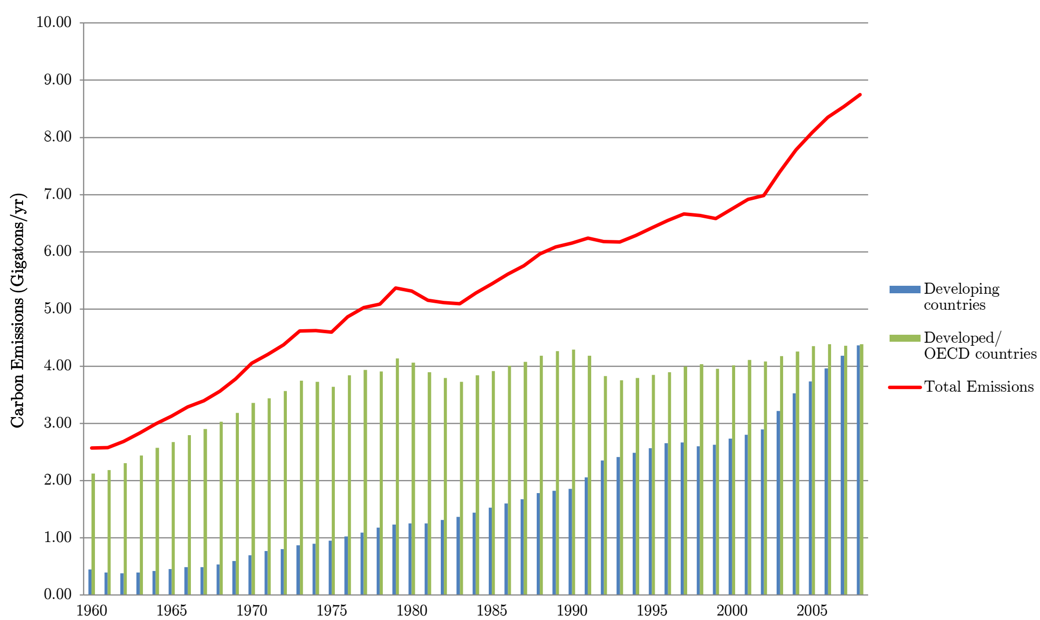


Figure 14 – *Developing Country CO_2 Emissions per capita: 122 Countries between 1980 – 2007*

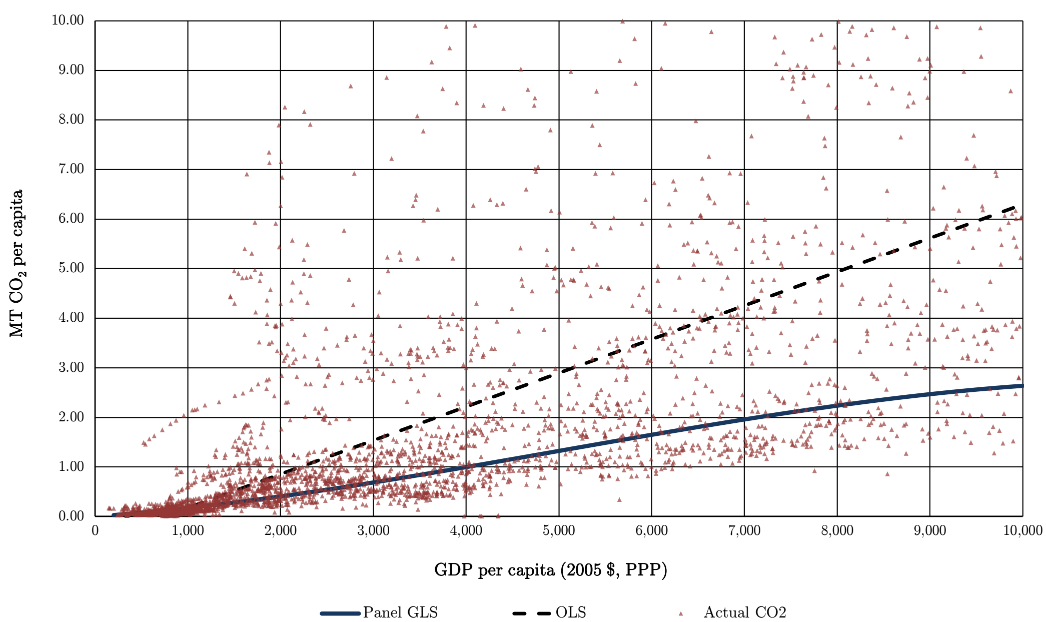


Figure 15 – *SES vs CO_2 Emissions in Developing Countries*

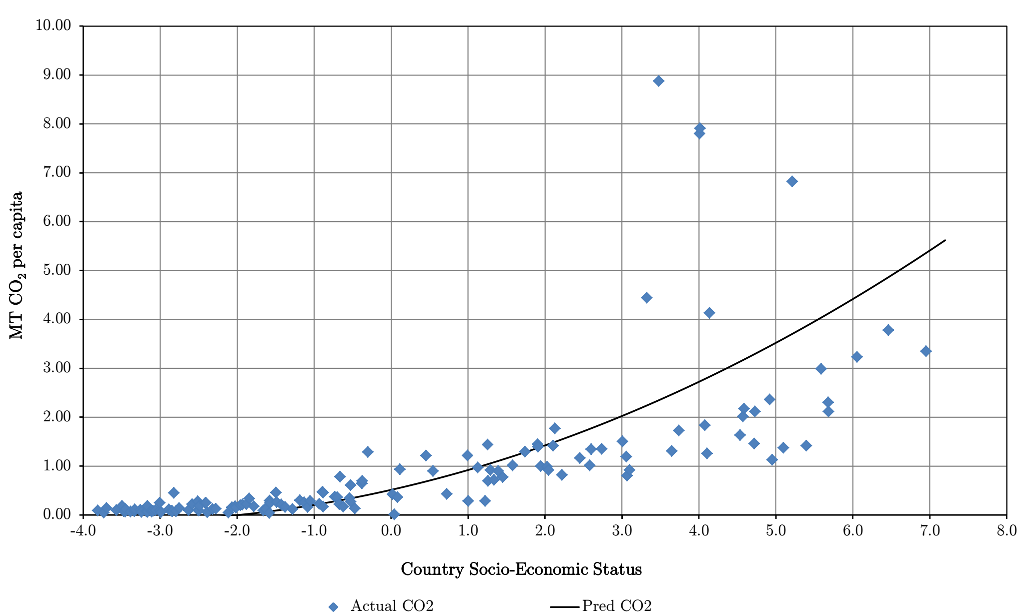


Figure 16 – *Comparing SES and GDP per capita predictors of per capita CO_2 Emissions in Developing Countries*

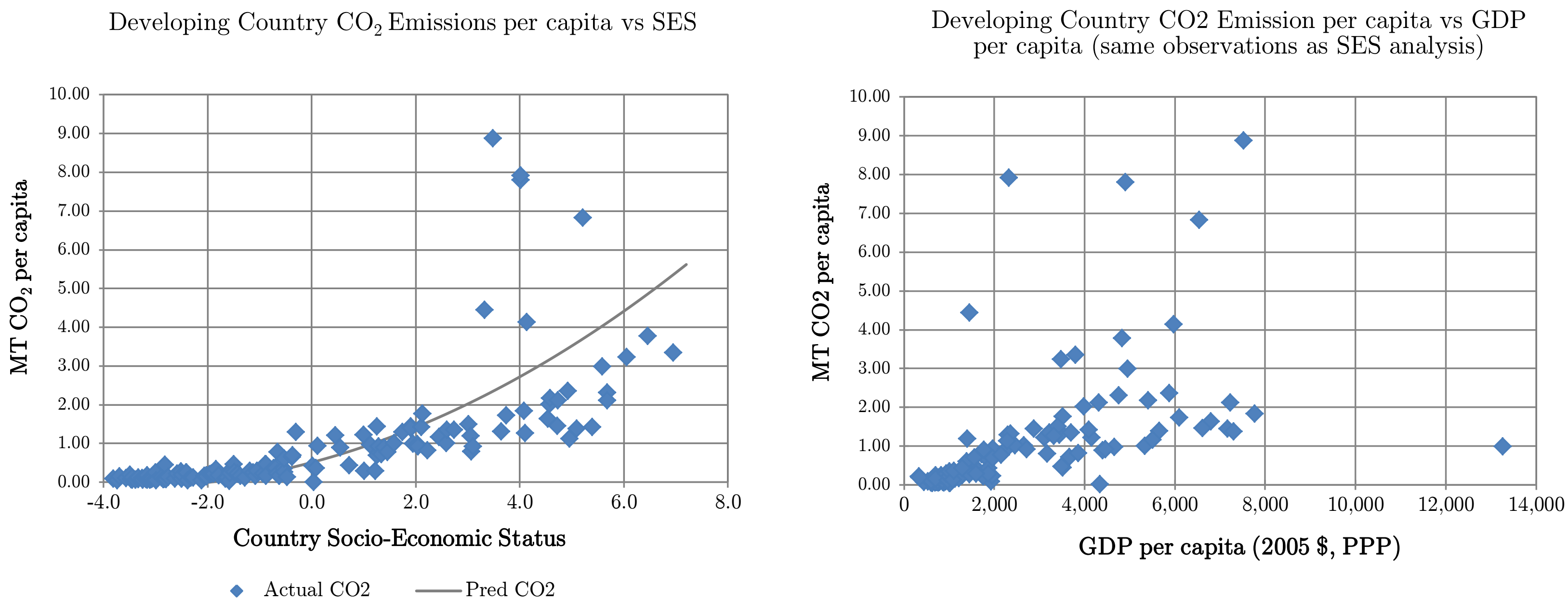
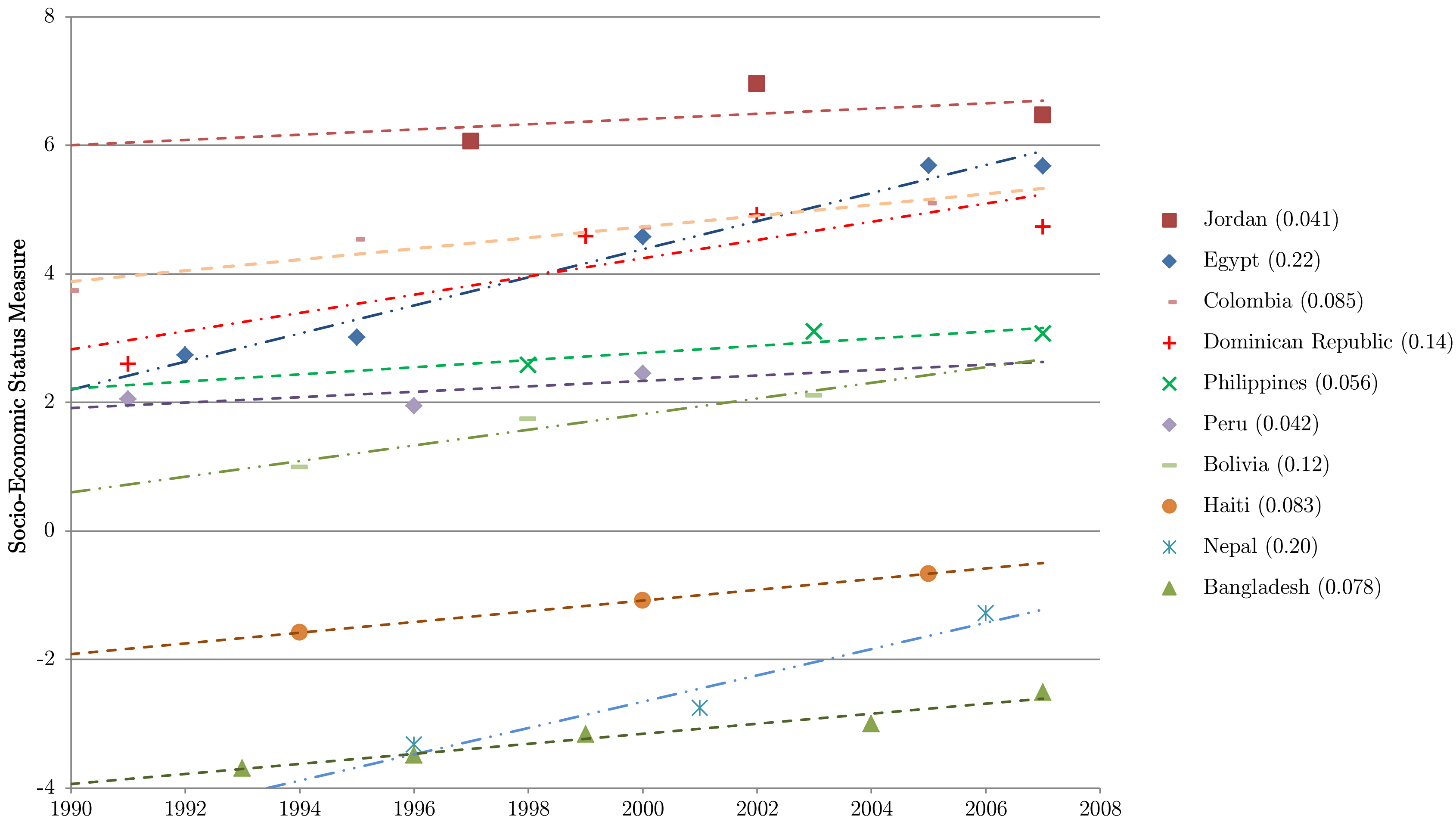
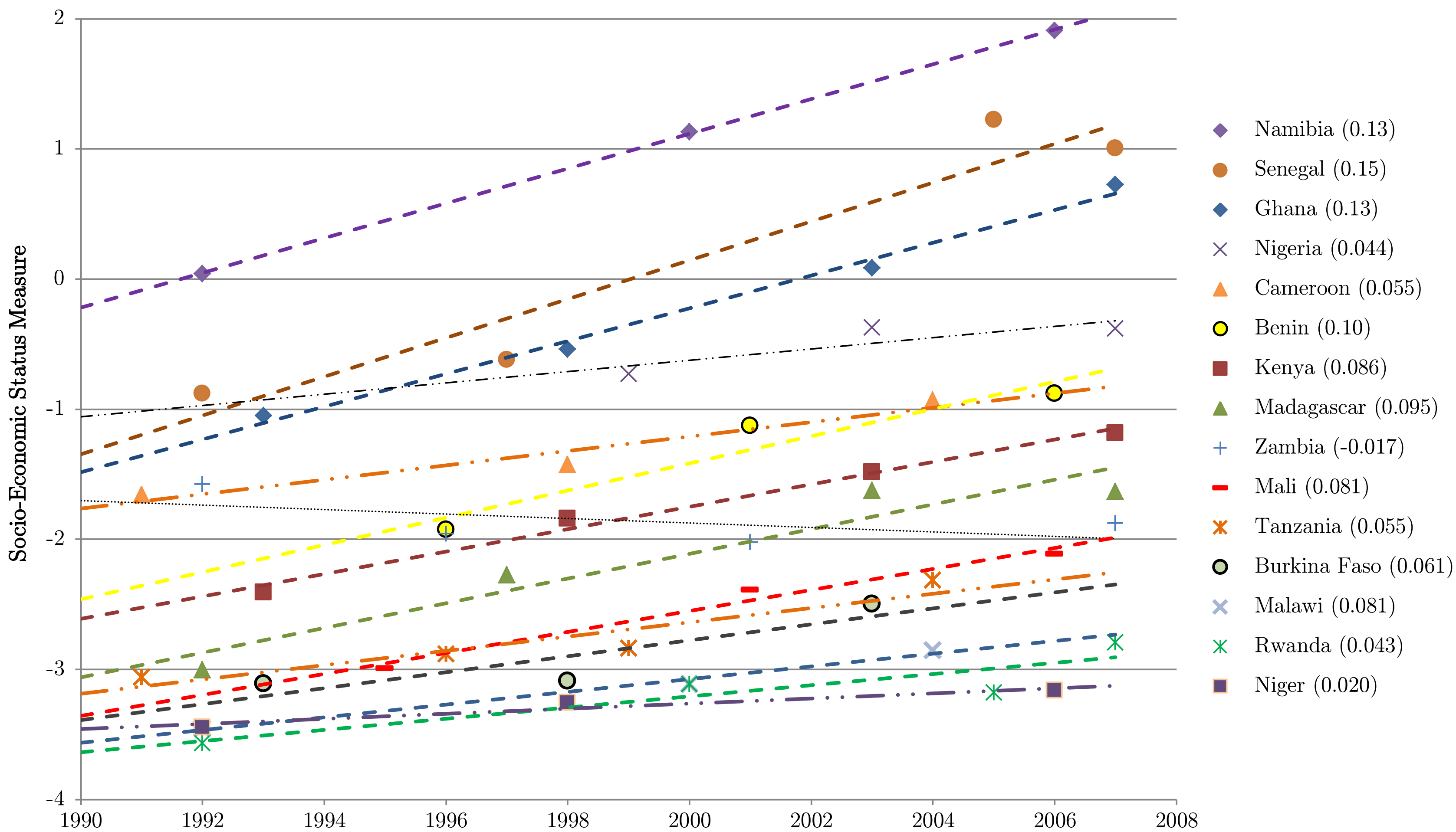


Figure 17 – *SES Trends for non–African Countries*



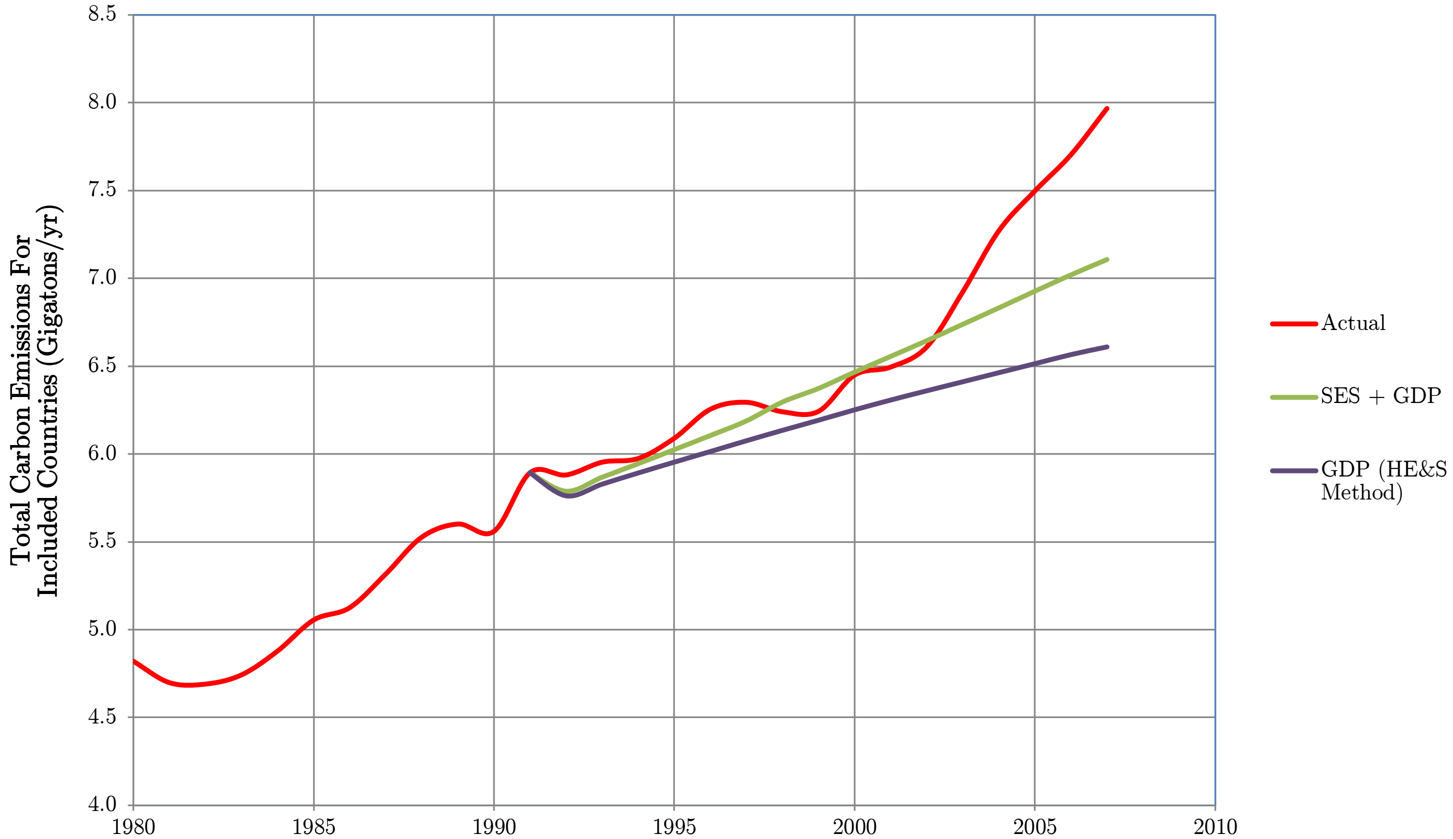
The slope of the trend line for each country is displayed in parentheses in the legend.

Figure 18 – *SES Trends for African Countries*



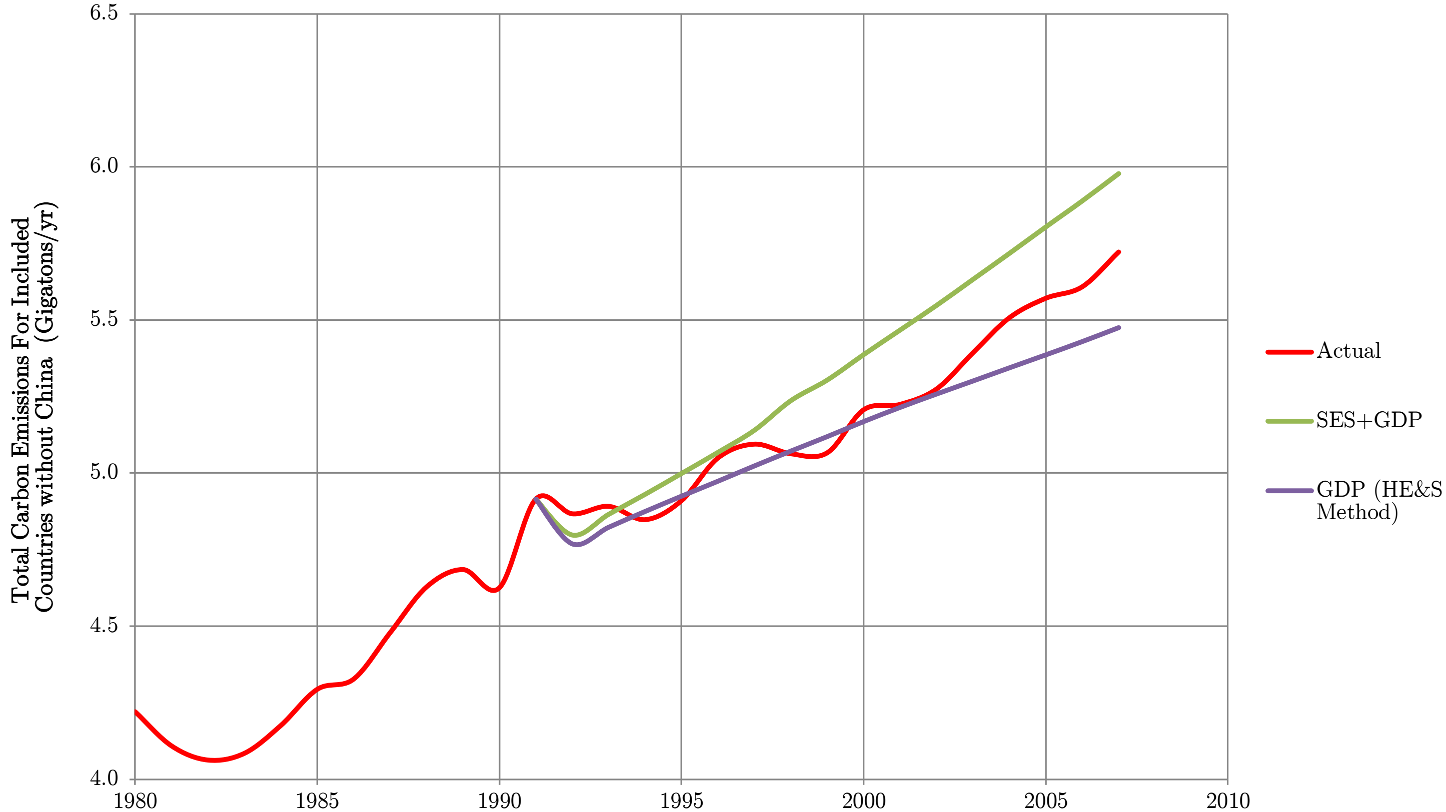
The slope of the trend line for each country is displayed in parentheses in the legend.

Figure 19 – *Comparing Actual Carbon Emissions to in-sample Predictions using the SES EKC Combined with the GDP EKC (SES + GDP), as well as the GDP EKC on its own (GDP)*



Note: Minor differences in total emissions between this figure and Figures 11 and 12 are the result of the exclusion of some smaller countries that do not have consistently reported data between 1980 and 2008 in this comparison.

Figure 20 – *Comparing Actual Carbon Emissions Excluding China to in-sample Predictions*



Appendix H: Tables

Table 38 – *Mean Household Percentages and Component Weights from SES PCA*

	Mean	**PC_1 Weight**		**Mean**	**PC_1 Weight**
Piped Water	41.8	0.269	Electricity	45.7	0.323
Well Water	33.7	−0.225	Radio	58.1	0.201
Surface Water	16.4	−0.177	Television	36.9	0.323
Rain Water	1.6	0.078	Telephone	11.8	0.274
Tanker Truck	1.1	0.049	Refrigerator	24.4	0.328
Bottled Water	2.5	0.124	Bicycle	22.4	−0.225
			Motorcycle	7.1	0.052
			Car	7.6	0.259
Flush Toilet	27.1	0.307	Natural Floor	43.0	−0.295
Pit Toilet	42.6	−0.147	Rudimentary	9.6	0.099
No Facility	28.2	−0.192	Finished	46.8	0.248
Countries			63		
PC_1 λ/n			0.42		

Table 39 – *Developing Countries Included in the Demographic and Health Surveys for which Household Characteristic and Possession Data is Available*

Angola	Egypt	Moldova	Senegal
Armenia	Eritrea	Madagascar	Sierra Leone
Azerbaijan	Ethiopia	Maldives	Swaziland
Benin	Gabon	Mali	Togo
Burkina Faso	Ghana	Mozambique	Turkmenistan
Bangladesh	Guinea	Mauritania	Tanzania
Bolivia	Guatemala	Malawi	Uganda
Brazil	Haiti	Namibia	Ukraine
Central African Rep	Indonesia	Niger	Uzbekistan
Cote d'Ivoire	India	Nigeria	Vietnam
Cameroon	Jordan	Nicaragua	Yemen
Congo	Kenya	Nepal	South Africa
Congo Demo Rep	Kyrgyz Republic	Pakistan	Zambia
Colombia	Cambodia	Peru	Kazakhstan
Comoros	Liberia	Philippines	Chad
Dominican Rep	Morocco	Rwanda	Lesotho

Table 40 – *Per Capita CO_2 Emissions Pooled OLS and GLS Fixed Effects Regression Results for the Full Sample of Developing and Developed Countries*

	——OLS——	——GLS, FE——
GDP (2005 $/person, PPP)	0.677***	0.732***
	(0.028)	(0.073)
GDP^2	−0.014***	−0.016***
	(0.0015)	(0.0022)
GDP^3	1.0e-04***	1.0e-04***
	(2.0e-05)	(2.0e-05)
Trend	−0.031***	−0.027***
	(0.006)	(0.007)
Constant	0.094	0.020
	(0.105)	(0.338)
Observations	4,190	4,190
R^2	0.651	0.330
F	1,162	45.91***
Hausman Test of GDP Exogeneity	107	

Robust standard errors clustered at the country level in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 41 – *Per Capita CO_2 Emissions pooled OLS Regression Results using the SES Measure*

	——OLS——
SES	0.356***
	(0.089)
SES^2	0.049***
	(0.015)
SES^3	-0.006
	(0.004)
Trend	0.005
	(0.010)
Constant	0.513***
	(0.141)
Observations	140
R^2	0.51
$F(4, 135)$	74.3 ***

Robust standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 42 – *Results from Estimation of per capita GDP Growth Rate Model*

$ln(y_{it+1}) - ln(y_{it})$	——GLS, FE——
$ln(y_{it})$	0.0329*
	(0.0155)
$[ln(y_{it})]^2$	−0.0018*
	(9.0e-04)
Constant	−0.134*
	(0.0657)
Observations	4,190
R^2	0.06
$F(2, 4187)$	7.73***

Robust standard errors clustered at the country level in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

Aldy, J.E. An environmental Kuznets curve analysis of US state-level carbon dioxide emissions. *Journal of Environment and Development*, 14:48–72, 2005.

Andresen, J.A.; Alagarswamy, G.; Rotz, C.A., and Lebaron, A.W. Weather impacts on maize, soybean, and alfalfa production in the Great Lakes region, 1895-1996. *Agronomy Journal*, 93(5):1059–1070, 2001.

Azomahou, T.; Laisney, F., and P.N., Van. Economic development and CO_2 emissions: A nonparametric panel approach. *Journal of Public Economics*, 90:1347–1363, 2006.

Baier, W. *Crop-weather models and their use in yield assessments*. Technical note. Secretariat of the World Meteorological Organization, 1977. ISBN 9789263104588.

Beck, Robert E., editor. *Waters and Water Rights*. Michie, 2004.

Bertinelli, L. and Strobl, E. The environmental Kuznets curve semi-parametrically revisited. *Economics Letters*, 88:350–357, 2005.

Blackburne, Edward F. and Frank, Mark W. Estimation of nonstationary heterogeneous panels. *The Stata Journal*, 7(2):197–208, 2007.

Cameron, A. Colin and Trivedi, Pravin K. *Microeconometrics Using Stata*. Stata Press, College Station, Texas, 2009.

Cole, M.A.; Rayner, A.J., and Bates, J.M. The environmental Kuznets curve: an empirical analysis. *Environment and Development Economics*, 2:401–416, 1997.

Dahl, Carol and Duggan, Thomas E. U.S. energy product supply elasticities: A survey and application to the U.S. oil market. *Resource and Energy Economics*, 18(4):243–263, 1996.

Dahl, Carol and Yucel, Mine. Testing alternative hypotheses of oil producer behavior. *Energy Journal*, 12(4):117–138, 1991.

Dasgupta, Partha S. and Heal, Geoffrey M. *Economic Theory and Exhaustible Resources*. James Nisbet and Co. and Cambridge University Press, Welwyn, England, 1979.

Dellapenna, Joseph W. Adapting riparian rights to the twenty-first century. *West Virginia Law Review*, 106:539–593, 2004.

Dijkgraaf, E. and Vollebergh, H.R.J. A test for parameter homogeneity in CO_2 panel EKC estimations. *Environmental and Resource Economics*, 32:229–239, 2005.

Dixon, Bruce L.; Hollinger, Steven E.; Garcia, Philip, and Tirupattur, Viswanath. Estimating corn yield response models to predict impacts of climate change. *Journal of Agricultural and Resource Economics*, 19(01):58–68, July 1994.

Draper, N.R. and Smith, H. *Applied Regressions analysis 2nd Edition*. John Wiley and Sons, New York, 1981.

Duchon, Claude E. Corn yield prediction using climatology. *Journal of Climate and Applied Meteorology*, 25(5):581–590, 1986.

Eheart, J. Wayland, editor. *Riparian Water Regulations: Guidelines for Withdrawal Limitations and Permit Trading.* American Society of Civil Engineers, 2002.

FAPRI, . Crop insurance: Background statistics on participation and results. Technical Report 10-10, Food and Agricultural Policy Research Institute at the University of Missouri, 2010.

Fekedulegn, B. Desta.; Colbert, J.J.; Hicks, R.R., and Schuckers, Michael E. Coping with multicollinearity: an example on application of principal components regression in dendroecology. Research Paper NE-721, U.S. Dept. of Agriculture, Forest Service, Northeastern Research Station, 2002.

Frankel, J.A. and Rose, A.K. Is trade good or bad for the environment? Sorting out causality. *The Review of Economics and Statistics*, 87(1):85–91, 2005.

Friedl, B. and Getzner, M. Determinants of CO_2 emissions in a small open economy. *Ecological Economics*, 45:133–148, 2003.

Galeotti, M. and Lanza, A. Richer and cleaner? A study on carbon dioxide emissions in developing countries. *Energy Policy*, 27:565–573, 1999.

Gamponia, Villmor and Mendelsohn, Robert. The taxation of exhaustible resources. *The Quarterly Journal of Economics*, 100(1):165–181, 1985.

Garcia, Philip; Offutt, Susan E.; Pinar, Musa, and Changnon, Stanley A. Corn yield behavior: Effects of technological advance and weather conditions. *Journal of Applied Meteorology*, 26(9):1092–1102, 1987.

Gaudet, Gérard. Natural resource economics under the rule of Hotelling. *Canadian Journal of Economics*, 40(4):1033–1359, 2007.

Goldberger, Arthur S. *A Course in Econometrics.* Harvard University Press, Cambridge, MA, 1991.

Goodwin, B.K. and Ker, A.P. Nonparametric estimation of crop yield distributions: Implications for rating group risk (RGR) crop insurance contracts. *American Journal of Agricultural Economics*, 80(1):139–153, 1998.

Griffin, James M. OPEC behavior: A test of alternative hypotheses. *American Economic Review*, 75(5):954–963, 1985.

Grossman, G.M. and Krueger, A.B. Economic growth and the environment. *Quarterly Journal of Economics*, 60:353–377, 1995.

Gwatkin, D.R.; Rutstein, S.; Johnson, K.; Suliman, E.; Wagstaff, A., and Amouzou, A. Socio-economic differences in health, nutrition, and population within developing countries: an overview. *Nigerian Journal of Clinical Practice*, 10(4):272–282, 2007.

Harri, Ardian; Erdem, Cumhur; Coble, Keith H., and Knight, Thomas O. Crop yield distributions: A reconciliation of previous research and statistical tests for normality. *Review of Agricultural Economics*, 31(1):163–182, 2009.

Hausman, Jerry A. and Taylor, William E. Panel data and unobservable individual effects. *Econometrica*, 49(6):1377–1398, 1981.

Hogan, William W. *World Oil Price Projections: A sensativity analysis*. Energy Environmental Policy Center, John F. Kennedy School of Government, Harvard University, Cambridge, MA, 1989.

Holtz-Eakin, D. and Selden, T.M. Stoking the fires? CO_2 emissions and economic growth. *Journal of Public Economics*, 57:85–101, 1995.

Hook, James E. Using crop models to plan water withdrawals for irrigation in drought years. *Agricultural Systems*, 45(3):271–289, 1994.

Houweling, T.A.J.; Kunst, A.E., and Mackenbach, J.P. Measuring health inequality among children in developing countries: does the choice of indicator of economic status matter? *International Journal for Equity in Health*, 2:8–20, 2003.

Hu, Qi and Buyanovsky, Gregory. Climate effects on corn yield in Missouri. *Journal of Applied Meteorology*, 42(11):1626–1635, 2003.

IPCC, . *Climate Change: The IPCC Scientific Assessment*. Cambridge University Press, Cambridge, United Kingdom, 1990.

IPCC, . *Climate Change 1992. The Supplementary Report to the IPCC Scientific Assessment*. Cambridge University Press, Cambridge, United Kingdom, 1992. Houghton, J.T., B.A. Callander and S.K. Varney (eds.).

IPCC, . *Special Report on Emissions Scenarios*. IPCC, Geneva, Switzerland, 2000. Nebojsa Nakicenovic and Rob Swart (eds.).

IPCC, . *Climate Change 2001: Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom, 2001. Watson, R.T. (ed.).

IPCC, . *Climate Change 2007: Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland, 2007. Pachauri, R.K and Reisinger, A. (eds.).

Jones, C.A. and Kiniry, J.R. *CERES-Maize: A simulation Model of Maize Growth and Development*. Texas A&M University Press, 1986.

Jones, Clifton T. OPEC behaviour under falling prices: Implications for cartel stability. *Energy Journal*, 11(3):117–129, 1990.

Kandel, Eugene. Competition or collusion: the case of OPEC. W.E. Simon Graduate School of Business, University of Rochester, Rochester, NY, 1990.

Kaufmann, Robert K. and Snell, Seth E. A biophysical model of corn yield: Integrating climatic and social determinants. *American Journal of Agricultural Economics*, 79(1): 178–190, 1997.

Klein, Christine A.; Angelo, Mary Jane, and Hamann, Richard. Modernizing water law: The example of Florida. *Florida Law Reivew*, 61(3), 2009.

Krichene, Noureddine. A simultaneous equations model for world crude oil and natural gas markets. Working Paper WP/05/32, International Monetary Fund, 2005.

Kunce, Mitch; Gerking, Shelby; Morgan, William, and Maddux, Ryan. State taxation, exploration, and production in the us oil industry. *Journal of Regional Science*, 43(4): 749–770, 2003.

LaRouche, Roger. Audit report: Stripper oil well property royalty rate reduction program, BLM. Program Audit 01-I-297, US Department of the Interior Office of Inspector General, 2001.

Levin, Andrew; Lin, Chien-Fu, and Chu, Chia-Shang James. Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1):1–24, 2002.

List, J.A. and Gallet, C.A. The environmental Kuznets curve: does one size fit all? *Ecological Economics*, 31:409–423, 1999.

Loehlin, J.C. *Latent variable models: an introduction to factor, path, and structural analysis*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1998.

MacAvoy, Paul. *Crude Oil prices as determined by OPEC and market fundamentals*. Ballinger, Cambridge, MA, 1982.

Maloney, Frank E.; Plager, Sheldon J., and Baldwin, Fletcher N. *Water law and administration the Florida experience*. University of Florida Press, 1968.

Maloney, Frank E.; Ausness, Richard C., and Morris, J. Scott. *A Model Water Code*. University of Florida Press, 1972.

Mancke, Richard B. The longrun supply curve of crude oil produced in the United States. *The Antitrust Bulletin*, 15:727–756, 1970.

Manne, A.S. and Richels, R.G. *Buying greenhouse insurance: The economic costs of* CO_2 *emissions limits*. The MIT Press, Cambridge, MA, 1992.

Marcus, R.R. and Kiebzak, S. The role of water doctrines in enhancing opportunities for sustainable agriculture in Alabama. *Journal of the American Water Resources Association*, 44(6):1578–1590, 2008.

McKenzie, D.J. Measure inequality with asset indicators. Technical report, Bureau for Research and Economic Analysis of Development, Center for International Development, Harvard University, 2003. BREAD Working Paper No. 042.

McNider, R.T.; Christy, J.R., and Hairston, J.E. Bringing agriculture back to water - a solution for the 21st century. In *AMS Forum: Living with a Limited Water Supply*. American Meteorological Society Annual Meeting, Jan 2005.

Mendelsohn, Robert Owen and Dinar, Ariel. Climate, water, and agriculture. *Land Economics*, 79(3):328–341, 2003.

Mendelsohn, Robert Owen; Nordhaus, William D., and Shaw, Daigee. The impact of global warming on agriculture: A Ricardian analysis. *American Economic Review*, 84 (4):753–771, 1994.

Millimet, D.L.; List, J.A., and Stengos, T. The environmental Kuznets curve: Real progress or misspecified models? *The Review of Economics and Statistics*, 85(4):1038–1047, 2003.

Moomaw, W.R. and Unruh, G.C. Are environmental Kuznets curves misleading us? the case of of CO_2 emissions. *Environment and Development Economics*, 2(4):451–464, 1997.

Myers, R.H. *Classical and modern regression with applications.* Duxbury Press, Boston, 1986.

Nordhaus, W.D. and Yohe, G.W. *Future paths of energy and carbon dioxide emissions, in: Changing climate: Report of the carbon dioxide assessment committee, National Research Council.* National Academy Press, Washington, D.C., 1983.

Norwood, Bailey; Roberts, Matthew C., and Lusk, Jayson L. Ranking crop yield models using out-of-sample likelihood functions. *American Journal of Agricultural Economics*, 86(4):1032–1043, 2004.

Ozaki, Vitor; Goodwin, Barry Kent, and Shirota, Ricardo. Parametric and nonparametric statistical modelling of crop yield: implications for pricing crop insurance contracts. *Applied Economics*, 40(9):1151–1164, 2008.

Panayotou, T. Economic growth and the environment. Technical report, Center for International Development at Harvard University, 2000. Working Paper no. 56, Environment and Development Paper No.4.

Park, Hun Myoung. Linear regression models for panel data using SAS, Stata, LIMDEP, and SPSS. working paper. Technical report, The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University, 2009. http://www.indiana.edu/ statmath/stat/all/panel.

Paudel, K.P.; Limaye, A.S.; Hatch, L.U.; Cruise, J.F., and Musleh, F. Development of an optimal water allocation decision tool for the major crops suring the water deficit period in the southeast United States. *Natural Resource Modeling*, 18(3):281–306, 2005.

Pesaran, M. Hashem and Smith, Ron. Estimating long–run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1):79–113, 1995.

Ramirez, O.A. Estimation of a multivariate parametric model for simulating heteroskedastic, correlated, non-normal random variables: The case of corn-belt corn, soybeans and wheat yields. *American Journal of Agricultural Economics*, 79:191–205, 1997.

Ramirez, O.A.; Misra, S.K., and Field, J.E. Crop yield distribution revisited. *American Journal of Agricultural Economics*, 85:108–120, 2003.

Rao, Nirupama S. *Taxation and the extraction of exhaustible resources: evidence from California oil production*. PhD thesis, Massachusetts Institute of Technology, 2010.

Reilly, J.M.; Edmonds, J.A.; Gardner, R.H., and Brenkert, A.L. Uncertainty analysis of the IEA/ORAU CO_2 emissions model. *Energy Journal*, 8(3), 1987.

Reisner, M. *Cadillac desert: the American West and its disappearing water*. Penguin Books, 1993.

Runge, E.C.A. and Odell, R.T. The relation between precipitation, temperature and the yield of corn on the agronomy south farm, Urbana, Illinois. *Agronomy Journal*, 50(8): 448–454, 1958.

Schlenker, Wolfram and Roberts, Michael James. Nonlinear effects of weather on corn yields. *Review of Agricultural Economics*, 28(3):391–398, 2006.

Schlenker, Wolfram; Hanemann, W. Michael, and Fisher, Anthony C. The impact of global warming on U.S. agriculture: An econometric analysis of optimal growing conditions. *The Review of Economics and Statistics*, 88(1):113–125, 2006.

Selden, T. and Song, D. Environmental quality and development: Is there a Kuznets curve for air pollution emissions? *Journal of Environmental Economics and Management*, 27: 147–162, 1994.

Slade, Margaret E. Tax policy and the supply of exhaustible resources: theory and practice. *Land Economics*, 60(2):133–147, 1984.

Slade, Margaret E. and Thille, Henry. Whither hotelling: Tests of the theory of exhaustible resources. *Annual Review of Resource Economics*, 1:239–259, 2009.

Smith, J. Warren. The effect of weather upon the yield of corn. *Monthly Weather Review*, 42(2):78–92, 1914.

Stern, D.I.; Common, M.S., and Barbier, E.B. Economic growth and environmental degradation: the environmental Kuznets curve and sustainable development. *World Development*, 24(7):1151–1160, 1996.

Sun, J.W. The nature of CO_2 emissions Kuznets curve. *Energy Policy*, 27:691–694, 1999.

Swanson, Earl R. and Nyankori, James C. Influence of weather and technology on corn and soybean yield trends. *Agricultural Meteorology*, 20(4):327 – 342, 1979.

Sweeney, James L. Chapter 17 – Economic theory of depletable resources: An introduction. In Kneese, Allen V. and Sweeney, James L., editors, *Handbook of Natural Resource and Energy Economics*, volume 3, pages 759 – 854. Elsevier, New York, 1993.

Tannura, Michael A.; Irwin, Scott H., and Good, Darrel L. Weather, technology, and corn and soybean yields in the U.S. corn belt. Marketing and Outlook Research Reports 37501, University of Illinois at Urbana-Champaign, Department of Agricultural and Consumer Economics, 2008.

Thompson, Louis M. Weather and technology in the production of corn in the U.S. corn belt. *Agronomy Journal*, 61(3):453–456, 1968.

Thompson, Louis M. Weather variability, climatic change, and grain production. *Science*, 188(4188):535–541, 1975.

Thompson, Louis M. Climatic change, weather variability, and corn production. *Agronomy Journal*, 78(4):649–653, 1986.

Thompson, Louis M. Effects of changes in climate and weather variability on the yields of corn and soybeans. *Journal of Production Agriculture*, 31(1):20–27, 1988.

U.S. Energy Information Administration. Oil stripper wells accounted for over 16% of U.S. oil production in 2009, July 2011. URL `http://http://www.eia.gov/todayinenergy/detail.cfm?id=2270`.

Vyas, S. and Kumaranayake, L. Constructing socio-economic status indices: how to use principal component analysis. *Health Policy and Planning*, 21(6):459–468, 2006.

Wallace, H.A. Mathematical inquiry into the effect of weather on corn yield in the eight corn belt states. *Monthly Weather Review*, 48(8):439–446, 1920.

Westerlund, Joakim. Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, 69(6):709–748, 2007.

Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2002.

Yavapolkul, Navin. Environmental Kuznets curve: Empirical investigation using non-parametric approach. Working Paper, 2005.

Yucel, Mine K. Severance tax shifting: A quantitative analysis. *Resources and Energy*, 10 (3):265–275, 1988.